ECOM - 5151
AUGUST 1967
AD

AD 660322

CHANNEL CAPACITY AND CODING

By

Louis Dean Duncan

ATMOSPHERIC SCIENCES LABORATORY
WHITE SANDS MISSILE RANGE, NEW MEXICO

20050208201

Distribution of this report is unlimited.

ECOM

UNITED STATES ARMY ELECTRONICS COMMAND

CHANNEL CAPACITY AND CODING*

By

LOUIS D. DUNCAN

ECOM - 5151 August 1967

DA PROJECT 1L013001A91A

ATMOSPHERIC SCIENCES LABORATORY
WHITE SANDS MISSILE RANGE, NEW MEXICO

---

*Dissertation for the Degree of Doctor of Philosophy, New Mexico State University, University Park, N. M., July 1967.

ABSTRACT

A channel is defined to be a triple $S = \{X, \mu(.|x), (Y, Y')\}$ where X and Y are abstract sets, Y' is a σ-algebra of subsets of Y, and $\mu(.|x)$ is a probability measure on (Y, Y') for each $x \epsilon X$. X and Y are usually assumed to be either subsets of the real numbers or subsets of Euclidean n-space. The channel has additive noise, μ, if $\mu(A|x) = \mu(A-x)$ for all $x \epsilon X$, $A \epsilon Y'$. For a given channel S and a given real number $0 < \lambda < 1$ a λ-code of length n for S is a set of pairs $(x_1, D_1) \ldots (x_n, D_n)$ where $x_i \epsilon X$, $D_i \epsilon Y'$, $D_i \cap D_j = \emptyset$ if $i \neq j$ and $\mu(D_i|x_i) \geq 1-\lambda$ for $i = 1, 2, \ldots, n$. The supremum of the nonempty set of integers N such that S admits a λ-code of length N is denoted by $N(S, \lambda)$.

Let $Q^*(S)$ be the set of all probability measures defined on X. For $\pi \epsilon Q^*(S)$. Let γ^π be the measure defined on (Y, Y') by

$$\gamma^\pi(B) = \int \mu(B|x)\,d\pi$$

for all $B \epsilon Y'$. For $x \epsilon X$, let $f^\pi(y|x) = d\mu(.|x)/d\gamma^\pi$ (the Radon-Nikodym derivative). The capacity, C, of the channel S is defined to be $\sup \{C(\pi) | \pi \epsilon Q^*(S), \text{ support of } \pi \text{ finite}\}$, where

$$C(\pi) = \iint \log f^\pi(y|x)\ \mu(dy|x)\pi(dx).$$

The following improvement of Fano's theorem is proved in Chapter II.

Theorem: Let S be a channel with capacity C, and let $0 < \lambda < 1$ be given. Then for any $n > 0$, $0 < t < 1$

$$(1-\lambda) \log N(S^n, \lambda) \leq nC + \log \left(\frac{1+t}{t^\lambda}\right).$$

In Chapter III the connected λ-code is defined and investigated. This code, $\{x_i, D_i\}_{i=1}^N$, is defined to be a λ-code where each D_i is a $\mathcal{U}$-connected set; $\mathcal{U}$ is a topology on Y. The supremum of the set $\{N | S$ admits a connected λ-code of length $N\}$ is denoted by $N(S, \lambda, \mathcal{U})$. Most of the results of this chapter are for a channel of type A - a channel with additive noise μ which is absolutely continuous with respect to Lebesgue measure, γ; X is a closed interval; Y is the real numbers; and, $\mathcal{U}$ is the usual topology for the reals.

Let $f = d\mu/d\gamma$, the Radon-Nikodym derivative. f is a bell function if there exists y_0 such that f is increasing for all $y \leq y_0$ and f is decreasing for all $y \geq y_0$. In Chapter III two conditions for $N(S, \lambda) = N(S, \lambda, \mathcal{U})$ are obtained when S is a channel of type A and f is a bell function. One of these conditions is necessary; the other is sufficient.

A technique for delineating those measures which cannot affect the value of $N(S, \lambda, \mathcal{U})$ is investigated in Chapter IV.

Sufficient conditions for a channel to have finite capacity are investigated in Chapter V. The main result is:

Theorem: Let S be a channel with additive noise μ which is absolutely continuous with respect to Lebesgue measure, γ. If there exists a choice for the Radon-Nikodym derivative, $f = d\mu/d\gamma$, such that $\int g(y)d\gamma < \infty$ where $g(y) = \sup \{f(y|x) | x \epsilon X\}$ then the capacity of S is finite.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND PRELIMINARIES

Most of the terminology and notations which will be used herein are standard; however, to remove any possible ambiguity, much of it will be explained. For a set A, $a \epsilon A$ means that a is a member of A; the complement of A will be written A^c. It will sometimes be convenient to refer to a member of A as a point of A. All logarithms will be to the base e. The convention will be adopted throughout to define the expression 0 log 0 to be equal to 0. Integrals will always be in the sense of Lebesgue integration. The meaning of unexplained terminology or notation of integration theory or measure theory will be that of Halmos [4], and the meaning of any unexplained probability theory terminology or notation will be that of Loeve [6].

Information theory is one of the youngest branches of applied probability theory. Its conception can, with certainty, be considered to be the appearance in 1948 of the now classical work of Shannon [7]. From the very begining, information theory presented to mathematicians a whole new set of problems, including some very difficult ones. It is quite natural that early investigators, including Shannon, whose basic goal was to obtain practical results,

were not able to give enough attention to these mathematical difficulties. Consequently at many points of their investigations, they were compelled either to be satisfied with reasoning of an inconclusive nature or to limit the set of objects studied.

Investigations with the aim of setting information theory on a solid mathematical basis have begun to appear in recent years. However, in most of these endeavors finiteness conditions have been placed on certain sets in order to establish the desired results.

One of the most important entities considered in the mathematical study of information theory is the concept of a channel. A <u>channel</u> is defined to be the triple $S = \{X, \mu(.|x), (Y, Y')\}$ where X is an arbitrary set, (Y, Y') is a measurable space, and $\mu(.|x)$ is a probability measure on (Y, Y') for each $x \in X$. The set X is usually referred to as the <u>input</u> or <u>input space</u>; the set Y is referred to as the <u>output</u> or <u>output space</u>. If both X and Y contain only a finite number of elements, the channel is said to be discrete; if X has a finite number of elements but Y is infinite (either countable or uncountable), S is called semi-continuous (the term "semi-continuous" is of engineering origin).

In most of the literature it is assumed that the channel is either discrete or semi-continuous. Such restrictions will not be made in this dissertation; it will be assumed throughout that X and Y are arbitrary sets unless there is a specific statement to the contrary.

A channel operates as follows. The existence of a sender and receiver is assumed. The function of the sender is to choose a member of X and transmit it. Since it is not necessary to have a precise definition of the term transmit, a somewhat heuristic explanation is given. Transmission consists of choosing a point $x \epsilon X$ and associating with x a point $y \epsilon Y$ which, in general, depends on x. The points x and y will be called the <u>transmitted symbol</u> (transmitted signal, symbol sent, etc.) and the <u>received symbol</u> (received signal, received point, etc.), respectively. If a particular transmitted symbol always results in the same received symbol, the transmission may be considered as a function T: $X \rightarrow Y$. The more interesting case is that in which a given transmitted signal does not always result in the same received symbol. In this case the function T must be considered as a function of x and another variable, called the <u>noise</u>. The received variable is considered to be a chance variable, i.e., specific occurrences are governed by probability. The only property of the transmission known by the sender and receiver is that for each x sent the

probability that the received symbol is a member of $A \epsilon Y'$ is $\mu(A|x)$. The receiver is capable of scanning through any finite class of sets of Y' and determining which, if any, contains the received symbol. After making this determination, the receiver then tries to decide, based upon some type of logical analysis, which member of X was actually transmitted.

The technique which is usually employed by the receiver to decide which point was transmitted involves a predetermined decision scheme which is known to both the sender and the receiver. There are, of course, many ways by which this decision sheme can be defined. The one which has become a standard in studies of information theory is as follows: Let $0 \leq \lambda < 1$ be given; let $(x_1, D_1), \ldots, (x_n, D_n)$ be members of $X \times Y'$ having the properties that $D_i \cap D_j = \emptyset$ if $i \neq j$, and such that $\mu(D_i|x_i) \geq 1-\lambda$. The sender and receiver then agree to consider only those members $x_1, x_2, \ldots, x_n \epsilon X$. If x_k is transmitted, the receiver scans through the sets $D_1, \ldots, D_n$ and determines which, if any, contains the received symbol. If the received symbol is in D_m, the receiver concludes that x_m was transmitted; if none of these sets contain the received symbol, any decision may be made. The receiver's conclusion will be correct with probability $\geq 1-\lambda$.

<u>1.1 Definition</u>: A set of pairs $\{(x_1, D_1), \ldots, (x_n, D_n)\}$ having the properties described above is called a <u>λ-code of length n.</u>

A quantity which will be of considerable importance later, in fact that subject of most of the important theorems in information theory, is defined below.

1.2 Definition: Let S be a channel. Given $0 \leq \lambda < 1$, let $N(S, \lambda)$ denote the supremum (sup.) of the non-empty set of integers N such that S admits a λ-code of length N.

Given n channels $S_m = \{X_m, \mu_m(\cdot|x_m), (Y_m, Y'_m)\}$, $m = 1, 2, \ldots, n$, one can form the product channel, $S^{(n)} = S_1 \times \ldots \times S_n$ in a natural way. In fact this channel is defined by $S^{(n)} = \{X^{(n)}, \mu^{(n)}(\cdot|u), (Y^{(n)}, Y^{(n)\prime})\}$ where $X^{(n)} = X_1 \times \ldots \times X_n$; $(Y^{(n)}, Y^{(n)\prime})$ is the product of the measurable spaces (Y_m, Y'_m), $m = 1, 2, \ldots, n$, as defined by Halmos [4]; and if $u \epsilon X$, i.e., $u = (x_1, \ldots, x_n)$ with $x_m \epsilon X_m$), then $\mu^{(n)}(\cdot|u)$ denotes the product probability measure on $Y^{(n)\prime}$ defined by $\mu^{(n)}(B_1 \times \ldots \times B_n) = \mu(B_1|x_1) \ldots \mu(B_n|x_n)$ where $B_1 \epsilon Y'_1, \ldots, B_n \epsilon Y'_n$.

1.3 Definition: The channel $S^{(n)}$ defined above is called a memoryless channel of length n. If $S_1 = S_2 = \ldots = S_n = S$, one writes $S^{(n)} = S^n$ and calls S^n the memoryless channel of length n generated by S. Any channel may be regarded as a memoryless channel of length 1.

1.4 Remark: The definitions of $N(S^{(n)}, \lambda)$ and $N(S^n, \lambda)$ follows immediately from definitions 1.2 and 1.3.

Suppose there are N distinct points in X which the sender wishes to transmit in such a manner that, for a predetermined $0 \leq \lambda < 1$, the probability that the receiver will wrongly deduce which point was sent is $\leq \lambda$. The channel S cannot necessarily perform this function if $N > N(S, \lambda)$. However, if for fixed λ, $N(S^n, \lambda)$ becomes unbounded with n, the problem can be solved by choosing an n_o such that $N(S^{n_o}, \lambda) > N$, establishing a one-to-one correspondence between the N points and a properly chosen set of N members of X^{n_o}, and using the channel S^{n_o}.

This problem and two equivalent (according to Wolfowitz [11]) versions are listed below:

Form I : Given N and λ, how small an n will suffice?

Form II : Given n and λ, how big an N can be achieved?

Form III: Given n and N, how small a λ can be achieved?

A companion problem to the above problem (call it the first) is the (second) problem of constructing a code to implement the answer to the first problem. In fact, it might be reasonably thought that the first problem could not be solved without a solution of the second. This is not the case, and, at present, existing knowledge about the first problem considerably exceeds

the existing knowledge about the second problem.

One of the main objectives of this dissertation is to define a technique for coding (called a connected λ-code) and to investigate what conditions must be placed on the channel so that such a code will provide a solution to the second problem.

The solution of the first problem is usually referred to as the coding theorem and its converse. Precise statements of these theorems must be reserved for later, since they involve terminology which has not yet been introduced

The other main objective of this dissertation is to obtain an improvement of the known results for the converse of the coding theorem for the general channel. It will be seen later that this amounts to obtaining an improvement of the known results for an upper bound for $N(S^n, \lambda)$ where S is a general channel (the channel where both the input and the output are arbitrary sets).

Before continuing toward these objectives, some basic notation, definitions, and theorems will be listed for future reference. Since information theory employs many of the tools of probability theory and measure and integration theory this list a fortiori contains results from these disciplines.

In addition to probability statements about the received symbol, many of the important results entail probabilistic statements about the transmitted symbol. The following developments will indicate the importance and application of such statements as well as provide a rigorous foundation for their formulation.

1.5 Definition: Let $S' = \{X, \mu(.|x), (Y, Y')\}$ be a given channel. The set $Q^* = Q^*(S')$ is defined to be the collection of all entities

$$\pi = \{X_o, X_o', \pi\}$$

of the following kind: The set X_o (called the support of π) is any fixed subset of X, X_o' is a σ -algebra of subsets of X_o containing all sets of the form

$$\{x | x\epsilon X_o, \ \mu(B|x) < \alpha, B\epsilon Y', \alpha \text{ real}\}$$

i.e., such that the function $\mu(B|x)$ is measurable in x for fixed $B\epsilon Y'$. Finally π is a probability measure on the σ-algebra X_o'.

The apparent ambiguity in letting the Greek letter π both represent and be a member of the entity $\{X_o, X_o', \pi\}$ will cause no confusion in usage and will allow for simplicity in notation.

1.6 Definition: Let (Y, Y') be a measurable space, and let μ and γ be measures on Y'. Then μ is absolutely continuous with respect to γ (written $\mu << \gamma$) if $\gamma(A) = 0$ implies $\mu(A) = 0$; μ is singular with respect to γ ($\mu \perp \gamma$) if there exists a set $A \epsilon Y'$ such that $\mu(A) = 0$ and $\gamma(A^c) = 0$.

The following well-known theorem is stated for completeness.

1.7 Lebesgue Decomposition Theorem: Let μ and γ be measures defined on the measurable space (Y, Y'); then μ can be written uniquely as the sum of two measures μ_1 and μ_2 where $\mu_1 << \gamma$ and $\mu_2 \perp \gamma$.

1.8 Remark: Any set D such that $\mu(D) = \mu(Y)$ is called a support of μ. This will be written $D = \text{spt } \mu$. If $\mu \perp \gamma$ D can be chosen such that $\gamma(D) = 0$. Whenever $\mu = \mu_1 + \mu_2$ with $\mu_1 << \gamma$ and $\mu_2 \perp \gamma$ a support, D, can always be chosen for μ_i, $i = 1, 2$ such that $\mu_j(D) = 0$ $j \neq i$. Throughout this dissertation such a choice for a support will always be implied.

1.9 Remark: Let μ and γ be defined as above and let $\mu = \mu_1 + \mu_2$ where $\mu_1 << \gamma$ and $\mu_2 \perp \gamma$. Then there exists a measurable set D such that $\gamma(D) = 0$ and $\mu_2(D^c) = 0$. The set D is called a singular set of μ with respect to γ. Let $g(y) = d\mu_1/d\gamma$

(the R-ion-Nikodym derivative), then $g(y)$ is a finite-valued, non-negative, real-valued, γ-measurable function, unique up to a set of γ-measure 0, such that

$$\mu_1(B) = \int g(y)\,d\gamma, \text{ for } B\varepsilon Y'$$

By definition of D, one is allowed to assume that

$$0 < g(y) < \infty \text{ if } y\varepsilon D^c,$$

$$g(y) = \infty \text{ if } y\varepsilon D.$$

Throughout this dissertation such a determination of Radon-Nikodym derivatives will always be chosen.

The preceding developments provide a suitable background for the definition of a new set of probability measures on (Y, Y'). Let $\pi\varepsilon Q^*$. For $B\varepsilon Y'$ define

$$\gamma^\pi(B) = \int \mu(B|x)\pi(dx).$$

It is easy to see that γ^π is a probability measure on (Y, Y') for each $\pi\varepsilon Q^*$.

Let $f^\pi(y|x) = d\mu_1/d\gamma^\pi$ where $\mu_1(.|x)$ is the absolutely continuous component of $\mu(.|x)$ with respect to γ^π. Loeve [6] has shown that $f^\pi(y|x)$ can be chosen such that it is jointly measurable in $x\varepsilon X_0$, $y\varepsilon Y'$ relative to the σ - algebra $X_0' \times Y'$. By definition of the set D in Remark 1.9, it is clear that $f^\pi(y|x)$

can (and will) be chosen so that it also satisfies this joint measurability condition in addition to the requirement specified in the remark.

The measure γ^{π} and the functions $f^{\pi}(y|x)$ will now be used to define an important characteristic of the channel, the Capacity.

Let $\pi \epsilon Q^*$ be arbitrary but fixed. Define

$$C(\pi) = \iint \log f^{\pi}(y|x)\ \mu(dy|x)\ \pi(dx),$$

$$= \iint f^{\pi}(y|x)\ \log f^{\pi}(y|x)\ \gamma^{\pi}(dy)\pi(dx)$$

if $\mu(.|x) << \gamma^{\pi}$ for almost all $[\pi]$ $x \epsilon X$. Otherwise, define $C(\pi) = +\infty$.

If the support of π is countable, the following equivalent definition of $C(\pi)$ will be convenient. If $\mu(.|x) << \gamma^{\pi}$, define $C(x|\pi) = \int \log f^{\pi}(y|x)\ \mu(dy|x)$. It is observed that, in this case, $\mu(.|x) << \gamma^{\pi}$ for almost all $[\pi]$ x in the support of π. Hence,

$$C(\pi) = \sum_{x \epsilon \mathrm{Spt}\pi} \pi(x)\ C(x|\pi).$$

1.10 Definition: Let $Q = \{\pi \epsilon Q^* |$ support of π is finite$\}$. Let $C = \sup \{C(\pi)|\pi \epsilon Q\}$. The quantity C is called the capacity of the channel.

The following important and somewhat surprising result in due to Kemperman [5].

<u>1.11 Theorem</u>: Let $C^* = \sup\{C(\pi) \mid \pi \epsilon Q^*\}$. Then $C^* = C$.

The coding theorem and its converse can now be stated. The following statements of these theorems are those of Wolfowitz [11].

<u>1.12 Theorem</u>: (The coding Theorem): Let $0 < \lambda < 1$ be given. Then there exists a positive constant K such that for any $n > 0$

$$N(S^n, \lambda) \geq e^{nC - K\sqrt{n}} .$$

<u>1.13 Theorem</u>: (The strong Converse). Let $0 < \lambda < 1$ and $\varepsilon > 0$ be given. Then for any n sufficiently large

$$N(S^n, \lambda) \leq e^{n(C + \varepsilon)}.$$

<u>1.14 Theorem</u>: (The weak Converse). Given $0 < \lambda < 1$. Then for all $n > 0$

$$(1-\lambda) \log N(S^n, \lambda) \leq nC + \log 2.$$

Theorem 1.12 was conjectured for the discrete channel by Shannon [7] in 1948. The first proof was given by Feinstein [1] in 1954. Essentially different proofs were given in 1957 by Shannon [8] and Wolfowitz [9]. Shannon also conjectured

theorem 1.14 for the discrete channel. The strong converse is due to Wolfowitz [9].

Wolfowitz [10] has shown that theorem 1.12 is true for a semicontinuous channel by approximating the semicontinuous channel by a discrete channel. The proof of theorem 1.13 for the semicontinuous channel is also due to Wolfowitz [10]. The following stronger version is due to Kemperman [5].

1.15 Theorem: Let $0 < \lambda < 1$ be given. Then for any semicontinuous memoryless channel $\underset{\cdot}{S}$ there exists a constant $K > 0$ such that for any $n > 0$

$$N(S^n, \lambda) \leq e^{nC + K\sqrt{n}} .$$

Theorem 1.14 is due to Fano [2] who proved it for the general channel. An essentially different proof of this theorem has been given by Kemperman [5]. A different expression for the right-hand side of the formula given in theorem 1.14 will be obtained in Chapter II. For small λ this will give a much better result (in the since of a smaller upper bound) than the one given in 1.14. The following well known theorem which will be needed for the proof is listed for reference.

<u>1.16 Theorem</u>: Let S_m, $m = 1, \ldots, n$ be arbitrary channels of capacity C_m. Then the capacity of the product channel $S^{(n)}$ is $C_1 + \ldots + C_n$.

CHAPTER II

AN IMPROVEMENT OF FANO'S THEOREM

It is not unusual in the field of mathematics for one to conjecture an extension of a known result to a more general setting without being able to obtain a proof. Sometimes this conjecture remains an open problem for many years. This is the current status of both the coding theorem and the strong converse for the general channel.

In the formulation of a theorem the primary objective of which is to obtain an upper bound for some quantity, one usually attempts to establish as small an upper bound as possible. The author is unaware of any theorem which gives a better upper bound for $N(S, \lambda)$ for a general channel than theorem 1.14, Fano's theorem. The theorem presented below gives an improvement, for small λ, of Fano's result. The proof of theorem 2.1 is a modification of the proof of 1.14 given by Kemperman [5].

2.1 Theorem: Let $S = \{X, \mu(.|x), (Y, Y')\}$ be a given channel. Then given $0 < \lambda < 1$, $0 < t < 1$, one has, for each positive n

$$(1-\lambda) \log N(S^n, \lambda) \leq nC + \log \left(\frac{1+t}{t^\lambda}\right),$$

where C is the capacity of S.

Proof:

It follows from theorem 1.16 that the capacity of the channel S^n is nC; therefore, it suffices to prove that

$$(1-\lambda)\log N(S,\lambda) \leq C + \log\left(\frac{1+t}{t^{\lambda}}\right).$$

Let N be any positive integer such that $N \leq N(S, \lambda)$. Then there exists a λ-code of length N; call it $\{(x_i, D_i)\}_{i=1}^{N}$. Let $A = \{x_i | (x_i, D_i)$ is a member of this λ-code$\}$. (This notation could be simplified by assuming that x_i's were distinct members of X). For each $x \varepsilon A$ let $N(x) = \sum_{i=1}^{N} \sum_{x=x_i} 1$, and let $\pi(x) = N(x)/N$. It is easy to see that π is a probability measure on X with finite support. (More precisely $\pi \varepsilon Q(S)$.)

For each $B \varepsilon Y'$ define

$$\gamma^{\pi}(B) = \sum_{x \varepsilon A} \mu(B|x)\pi(x).$$

As has been pointed out γ^{π} is a probability measure on (Y, Y') and $\mu(.|x) << \gamma^{\pi}$ for all $x \varepsilon A$. For $\pi(x) > 0$ let

$$f^{\pi}(y|x) = d\mu(.|x)/d\gamma^{\pi}.$$

Now, by definition,

$$\sum_{x \varepsilon A} \pi(x) \int f^{\pi}(y|x) \log f^{\pi}(y|x)\, \gamma^{\pi}(dy) = C(\pi) \leq C.$$

Observe that

$$N^{-1}\sum_{i=1}^{N} f^{\pi}(y|x_i) = N^{-1}\sum_{x\in A} N(x)\; f^{\pi}(y|x)$$

$$= \sum_{x\in A} \pi(x)\; f^{\pi}(y|x) = 1.$$

Similary,

$$N^{-1}\sum_{i=1}^{N} \int f^{\pi}(y|x_i)\log f^{\pi}(y|x_i)\gamma^{\pi}(dy)$$

$$= \sum_{x\in A} \pi(x) \int f^{\pi}(y|x) \log f^{\pi}(y|x)\gamma^{\pi}(dy) \le C.$$

A new set of functions is defined on Y as follows. For each $1 \le i \le N$, define

$$h_i(y) = \begin{cases} N \text{ if } y\epsilon D_i \\ t \text{ if } y\epsilon D_i^c \end{cases}.$$

Then $\sum_{i=1}^{N} h_i(y) \le N(1+t)$ for each y.

The desired results will now be obtained by analyzing the terms of the following equation:

$$N^{-1}\int\sum_{i=1}^{N} f^{\pi}(y|x_i)\log\frac{h_i(y)}{f^{\pi}(y|x_i)}\gamma^{\pi}(dy)$$

$$+ N^{-1} \sum_{i=1}^{N} \int f^{\pi}(y|x_i) \log f^{\pi}(y|x_i)\gamma^{\pi}(dy)$$

$$= N^{-1} \sum_{i=1}^{N} \int f^{\pi}(y|x_i) \log h_i(y)\gamma^{\pi}(dy).$$

For reference this will be called equation (*).

For each $0 \leq i \leq N$

$$\int f^{\pi}(y|x_i) \log h_i(y)\gamma^{\pi}(dy) = \int \log h_i(y)\mu(dy|x_i)$$

$$= \mu(D_i|x_i) \log N + \mu(D_i^c|x_i) \log t.$$

$$\geq (1 - \lambda) \log N + \lambda \log t.$$

Therefore, $N^{-1} \sum_{i=1}^{N} \int f^{\pi}(y|x_i) \log h_i(y)\gamma^{\pi}(dy)$

$$\geq (1 - \lambda) \log N + \lambda \log t.$$

The function $\log Z$ is concave; hence,

$$N^{-1} \sum_{i=1}^{N} f^{\pi}(y|x_i) \log \frac{h_i(y)}{f^{\pi}(y|x_i)} \leq \log N^{-1} \sum_{i=1}^{N} h_i(y) \leq \log (1 + t).$$

Therefore, $N^{-1} \int \sum_{i=1}^{N} f^{*}(y|x_i) \log \frac{h_i(y)}{f^{*}(y|x_i)} \gamma^{*}(dy) \leq \log (1 + t).$

Inserting these inequalitites into equation (*), one obtains

$$\log (1 + t) + C \geq (1 - \lambda) \log N + \lambda \log t,$$

which is equivalent to

$$(1 - \lambda) \log N \leq C + \log \left(\frac{1 + t}{t^{\lambda}}\right)$$

Since $N \leq N(S, \lambda)$ was arbitrary, the theorem is proved.

2.2 Remark: Given $0 < \lambda < 1/2$ let $f(t) = (1 + t)/t^{\lambda}$ for $t \in (0, 1)$; $f(t) = +\infty$ if $t \notin (0, 1)$. Then f has a minimum at $t = \lambda/(1 - \lambda)$. The function $g(\lambda) = \lambda^{-\lambda}(1 - \lambda)^{\lambda-1}$ is an increasing function of λ and $g(y) < 2$.

2.3 Corollary: Let S be a given channel. Then for $0 < \lambda < 1$ one has, for each positive n, $(1 - \lambda) \log N(S^{n}, \lambda) \leq nC - \lambda \log \lambda + \log(1 + \lambda)$. If $0 < \lambda < 1/2$ then

$$(1 - \lambda) \log N(S^{n}, \lambda) \leq nC - \lambda \log \lambda - (1 - \lambda) \log (1 - \lambda).$$

CHAPTER III

CONNECTED λ-CODES

Although for a given channel, S, the value of $N(S, \lambda)$ may be known for each $0<\lambda< 1$, the actual construction of a λ-code of length $N(S, \lambda)$ may be quite difficult. In addition, for practical reasons, it may be desirable to place various restrictions upon the entities of the code. Such restrictions may, of course, preclude the possibility of attaining a code of maximal length. In this chapter, a λ-code with a specified restriction is defined and analyzed.

In the operation of a channel, one of the functions of the receiver is to scan through the sets D_i of a given λ-code $\{x_i, D_i\}_{i=1}^{N}$ and determine which of these, if any, contains the received signal. This determination may be quite difficult, the degree of the difficulty depending on the nature of the sets $\{D_i\}_{i=1}^{N}$. (Recall that the only restriction placed on the D_i's is that they may be members of Y'.)

<u>3.1 Definition</u>: Let $\mathcal{U}$ be a topology on Y. A <u>$\mathcal{U}$-connected λ-code</u> is a λ-code $(x_1, D_1), \ldots, (x_n, D_n)$ where D_i is a $\mathcal{U}$-connected set for each $i=1, 2, \ldots, n$.

<u>3.2 Definition</u>: For a given channel and a fixed real number $0<\lambda <1$, let $N(S, \lambda, \mathcal{U})$ denote the supremum of the

non-empty set of integers N such that S admits a $\mathcal{U}$-connected λ-code of length N.

It is clear that $N(S, \lambda, \mathcal{U}) \leq N(S, \lambda)$ for each $0 < \lambda < 1$ and any topology $\mathcal{U}$. It is also clear that equality need not hold unless some restrictions are imposed upon the channel. It appears intuitively clear that these restrictions should be placed on the set of probability measures $\{\mu(.|x) | x \varepsilon X\}$ and must include restrictions upon the supports of these measures. Although it would be desirable to let X and Y be arbitrary sets and $\mathcal{U}$ be any topology on Y, such generality leads to complicated and unwiedly analysis and yields very few results. In this chapter some restrictions will be placed upon the input space and the output space as well as the set of probability measures. These restrictions will allow for almost all of the practical physical situations one might expect to encounter.

The following example indicates some of the restrictions which must be placed upon this set of probability measures.

<u>3.3 Example:</u> Let $X = \{0, 1\}$; let $Y = [0, 1]$; let Y' be the Borel sets on $[0, 1]$; and let γ be Lebesgue measure. The measures $\mu(.|1)$ and $\mu(.|0)$ are defined as follows: For $B \varepsilon Y'$ define

$$\mu(B|0) = \int_{B_1} (-4x+2)d\gamma + \int_{B_2} (4x-2)d\gamma,$$

$$\mu(B|1) = \int_{B_1} 4x\,d\gamma + \int_{B_2} (-4x+4)\,d\gamma,$$

where $B_1 = B \cap [0, 1/2]$ and $B_2 = B \cap [1/2, 1]$.

Let U be the usual topology for [0, 1]. Since X contains only 2 elements $N(S, \lambda) \leq 2$ for all $0 < \lambda < 1/2$. Observe that $\mu([0, 1/4] \cup [3/4, 1]|0) = 3/4 = \mu([1/4, 3/4]|1]$. Hence $N(S, 1/4) \geq 2$. It is easy to see that if I_o is any interval contained in Y such that $\mu(I_o|0) \geq 3/4$ then $\gamma(I_o) > 1/2$; also if I_1 is any interval contained in Y such that $\mu(I_1|1) \geq 3/4$ then $\gamma(I_1) \geq 1/2$. It follows that $N(S, 1/4, U) = 1$.

In the example above, one is able to obtain a longer non-connected code becase of the nature of the Radon-Nikodym derivative of $\mu(.|0)$ with respect to Lebesgue Measure. It will be observed that if one chooses $0 < \lambda < 1$ and considers any connected set A such that $\mu(A|0) = 1-\lambda$, then there exists a set B such that $\mu(B|0) = 1-\lambda$ and $\gamma(B) < \gamma(A)$.

Although it was not difficult to show that $N(S, 1/4) = 2$ while $N(S, 1/4, U) = 1$, it is easy to see that this problem could rapidly become difficult as the number of points in X is increased. Given $x \epsilon X$, let $f(y|x) = d\mu(.|x)/d\gamma$. The relationship between $N(S, \lambda)$ and $N(S, \lambda, U)$ for arbitrary λ is very difficult, in fact, almost impossible, to determine when X is infinite, unless one requires

some type of uniformity among these functions. Such a restriction, to be rigorously defined later, will be placed upon the channels investigated in this chapter.

The following auxiliary results will be needed later. Throughout the remainder of this chapter R will denote the real numbers. $\mathcal{B}$ will be the Borel sets, γ will be Lebesgue measure, and $\mathcal{U}$ will be the usual topology for the reals. If Y is a subset of R then the σ-algebra Y' will be the Borel sets on Y.

3.4 Theorem: Let μ be a totally finite measure of $(R, \mathcal{B})$. Let ψ be any measure defined an $(R, \mathcal{B})$. Let $\alpha = \mu(R)$, let $0 < \lambda \leq \alpha$ be given, and let $\mathcal{A} = \{A\varepsilon\mathcal{B}: \mu(A) \geq \alpha-\lambda\}$. Then $\mathcal{A}$ contains a member of minimal ψ-measure.

Proof:

Let $\mu = \mu_1 + \mu_2$ where $\mu_1 \ll \psi$ and $\mu_2 \perp \psi$. Let $A_0 = \text{spt}\ \mu_2$. If $\mu(A_0) \geq \alpha-\lambda$ the theorem is trivially true. Suppose not. Let $\beta = \alpha-\lambda-\mu(A_0)$. Let $f(y) = d\mu_1/d\psi$. Let $E_z = \{y: f(y) \geq z\}$. Observe $z_1 < z_2$ implies $E_{z_1} \supset E_{z_2}$. Note that E_z is a ψ-measurable set for each $z\varepsilon R$. Let $I(z) = \int_{E_z} f(y)\, d\psi$. Then $I(z)$ is non-negative, monotone non-increasing and left-continuous. Let $z_0 = \inf\{z: I(z) \leq \beta\}$. Observe that $I(z_0) \geq \beta, z_0 > 0$.

Now if $B \subset E_{z_o}$ and $C \subset E^c_{z_o}$, then $\int_C f(y)d\psi \geq \int_B f(y)d\psi > 0$ implies $\psi(C) > \psi(B)$. [If $y \in C$ then $f(y) < z_o$ while for $y \in B$, $f(y) \geq z_o$]. Hence if $I(z_o) = \beta$ then $E_{z_o} \cup A_o$ is clearly a set of minimal ψ-measure such that $\mu(E_{z_o} \cup A_o) \geq \alpha-\lambda$. Suppose $I(z_o) > \beta$. Let $E' = \{y: f(y) > z_o\}$ and $E'' = \{y: f(y) = z_o\}$. If $\mu(E') = \beta$ then $A_o \cup E'$ is the required set. If $\mu(E') < \beta$ then $\mu(E'') > \beta - \mu(E') > o$. Let E''' be a subset of E'' of ψ-measure $(\beta - \mu(E'))/z_o$. Then $\mu(E' \cup E''') = \beta$ and $E' \cup E''$ has minimal ψ-measure. Hence $A = A_o \cup E' \cup E'''$ is a set such that $\mu(A) \geq \alpha-\lambda$ and such that $\psi(B) \geq \psi(A)$ for any set B such $\mu(B) \geq \alpha-\lambda$.

3.5 Remark: Let A be any set of minimal ψ-measure such that $\mu(A) \geq \alpha-\lambda$. Let B = spt μ_2 and let $y_o = \inf\{y \mid \mu_1(-\infty, y) \geq \alpha-\lambda - \mu_2(B)\}$. Let $f(y) = d\mu_1/d\psi$. If $f(y_1) \geq f(y_2)$ for almost all $[\psi]$ $y_1 < y_o < y_2$ then A can be chosen to be $A = B \cup (-\infty, y_o)$.

The proof of the following two Corollaries comes immediately from the proof of theorem 3.4.

3.6 Corollary: Let f be a non-negative real-valued Lebesgue measurable function. For $B \in \mathcal{B}$ and $0 \leq \alpha < \int_B f(y)d\gamma$ there exists a set $A \subset B$ of minimal γ-measure such that $\int_A f d\gamma \geq \alpha$.

3.7 Corollary: Let μ be a totally finite measure defined on $(R, \mathcal{B})$. Let $\alpha = \mu(R)$. If $\mu \ll \gamma$, then there exists $A \in \mathcal{B}$ such that $\mu(A) = \alpha-\lambda$ and $\gamma(B) \geq \gamma(A)$ for any $B \in \mathcal{B}$ such that $\gamma(B) \geq \alpha-\lambda$.

In several of the theorems developed in this chapter, it will be hypothesized that $\mu(.|x) << \gamma$ for all $x \epsilon X$. The following theorem gives a partial justification for such requirements.

3.8 Theorem: Let μ be a totally finite measure on $(R, \mathcal{B})$. Suppose that for each $0 \leq \lambda \leq \mu(R)$ there exists $A_\lambda \epsilon \mathcal{B}$ of minimal γ-measure such that $\mu(A_\lambda) \geq \mu(R)-\lambda$ and A_λ is connected. Let $\mu = \mu_1 + \mu_2$ where $\mu_1 << \gamma$ and $\mu_2 \perp \gamma$. Then there exists a support of μ_2 which contains at most one point ($|\text{spt } \mu_2| \leq 1$).

Proof:

Suppose $|\text{spt } \mu_2| \geq 2$. Let $\mathcal{A} = \{A | A = \text{spt } \mu_2$ and $\gamma(A) = 0\}$. Then given $A \epsilon \mathcal{A}$, A is not connected. Let $\lambda = \mu(R) - \mu_2(R)$ and let A_λ be any set of minimal γ-measure such that $\mu(A_\lambda) = \mu(R) - \lambda$. Observe that, given $A \epsilon \mathcal{A}$, $\mu(A) = \mu(R) - \lambda$. Hence $\gamma(A_\lambda) = 0$. Therefore $\mu(A_\lambda) = \mu_2(A_\lambda)$. It follows that $A_\lambda \epsilon \mathcal{A}$; hence is not connected.

3.9 Lemma: Let $f: R \to R$ be a.e. continuous and non-negative such that $0 < \int_R f d\gamma = \alpha < \infty$. Suppose there exists $x_1 < x_2 < x_3$ such that f is continuous at x_i $i=1, 2, 3$, and $f(x_1) > f(x_2) < f(x_3)$. Then there exists $\lambda > 0$ such that, if A is any set of minimal γ-measure such that $\int_A f d\gamma \geq \alpha-\lambda$, A is not connected.

Proof:

Let $E = \{x: f(x) \geq 1/2\, \{f(x_2) + \min (f(x_1), f(x_3))\}$. Let $\beta = \int_E f(x) d\gamma$,

Let $E_1 = E \cap \{x: x \leq x_2\}$,

$E_2 = E \cap \{x: x \geq x_2\}$,

and

$$\beta_i = \int_{E_i} f(x) d\gamma \quad i = 1, 2.$$

Since f is continuous at x_1, x_2 and x_3, it is easy to see that $\beta_1 \neq o \neq \beta_2$ and E is not connected.

Let $\lambda > o$ be chosen such that $\beta_1 + \beta_2 \geq \alpha - \lambda > \max (\beta_1, \beta_2)$.

Now $x \in E$ implies $f(x) \geq 1/2\, \{f(x_2), \min (f(x_1), f(x_3))\}$ and $x \in E^c$ implies that $f(x) < 1/2\, \{f(x_2), \min (f(x_1), f(x_3))\}$. Thus, if A is any set of minimal γ-measure such that $\int_A f(x) d\gamma \geq \alpha - \lambda$, then $\gamma(A \cap E^c) = 0$. Moreover, $\beta_1 > o$, $\beta_2 > o$, and $\alpha - \lambda > \max (\beta_1, \beta_2)$ implies $A \cap E_1 \neq \emptyset$, $A \cap E_2 \neq \emptyset$. Hence A is not connected.

With the aid of Lemma 3.9, one can characterize those a.e. continuous summable functions $f: R \to R$ such that given $o < \lambda < \int_R f d\gamma$ there exists a set A_λ, of minimal γ-measure for which $\int_{A_\lambda} f d\gamma \geq \int_R f d\gamma - \lambda$, which is connected.

3.10 Definition: A bell function is any function $f: R \to R$ such that there exists x_o such that f is monotone nondecreasing for $x < x_o$ and f is monotone nonincreasing for $x > x_o$.

3.11 Theorem: Let f be a non-negative a.e. continuous real-valued function defined an R such that $o < \int_R f d\gamma = \alpha < \infty$. Given $o < \lambda < \alpha$ let B be any set of minimal γ-measure such that $\int_B f d\gamma \geq \alpha - \lambda$. Then there exists for every λ a connected set A such that $\int_A f d\gamma = \int_B f d\gamma$ and $\gamma(A) = \gamma(B)$ iff f is a bell function a.e.

Proof:

Observe that if f is not a bell function a.e. then there exists points of continuity $x_1 < x_2 < x_3$ such that $f(x_1) > f(x_2) < f(x_3)$. Thus, the necessary part follows immediately from Lemma 3.9. The sufficient part will be proved by constructing the set A. The construction is similiar to that used in 3.4.

Assume f is a bell function. Let $E_y = \{x | f(x) \geq y\}$. Let $I(y) = \int_{E_y} f(x) d\gamma$. Let $y_o = \inf \{y: \ I(y) \leq \alpha - \lambda\}$. Then $I(y_o) \geq \alpha - \lambda$. Let $E'_{y_o} = \{x: \ f(x) > y_o\}$ and $E''_{y_o} = \{x: \ f(x) = y_o\}$. Now it is clear that if A_1 and A_2 are γ-measurable sets such that $A_1 \subset E_{y_o}$ and $A_2 \subset E^c_{y_o}$ and $o < \gamma(A_2) \leq \gamma(A_1)$ then $\int_{A_1} f d\gamma > \int_{A_2} f d\gamma$. Thus if $\gamma(E''_{y_o}) = o$ the proof is completed. Suppose $\gamma(E''_{y_o}) \neq o$. Then, $E''_{y_o} = [a_1, b_1] \cup [a_2, b_2]$ and $E'_{y_o} = (b_1, a_2)$ where $b_1 \leq a_2$. Clearly one can choose $a \in [a_1, b_1]$, $b \in [a_2, b_2]$

such that $[b_1 - a + b - a_2]y_0 = \omega - \lambda - \int_{E'_{y_0}} f\,d\gamma$. Now $A = [a, b]$ is the required interval.

The preceding results will be used in the analysis of an important special channel which will now be introduced.

Let S be a given channel. If anytime a value x is transmitted, the receiver is able to determine from the received symbol that x was the point set, then the channel is called noiseless. Such channels rarely occur in practice and are of little mathematical interest. On the other hand, if there is a definite positive probability that the receiver's decision will be wrong, then the channel is called noisy. An important type of noisey channel is the channel with additive noise.

3.12 Definition: Let $S = \{X, \mu(\cdot|x), (Y, Y')\}$ be a given channel. S has additive noise if there exists a probability measure μ on (Y, Y') such that $\mu(A|x) = \mu(A-x)$ for each $x \in X$, $A \in Y'$. The measure μ is called the noise.

3.13 Remark: In order to assure that the operations indicated in 3.12 are well defined, it will be assumed throughout the remainder of this dissertation that if S is a channel with additive noise then $(Y, +)$ is a group.

3.14 Theorem: Let S be a channel with additive noise μ. If $\mu \ll \gamma$, then $\mu(\cdot|x) \ll \gamma$ for each $x \in X$ and $f(y|x) = f(y-x)$

where $f(y) = d\mu/d\gamma$ and $f(y|x) = d\mu(.|x)/d\gamma$.

Proof:

By the Radon-Nikodym theorem

$$\mu(A|x) = \mu(A-x) = \int_{A-x} f(y)\,d\gamma = \int_A f(y-x)\,d\gamma.$$

Hence, by the absolute continuity of the integral, $\mu(A|x) << \gamma$. Also, by the Radon-Nikodym theorem

$$\mu(A|x) = \int_A f(y|x)\,d\gamma.$$

Therefore, $f(y|x) = f(y-x)$ for almost all $[\gamma]$ $y \in Y$.

Now it is clear that to determine a necessary and sufficient condition for $N(S, \lambda, \mathcal{U}) = N(S, \lambda)$, where S is a channel with additive noise, one need only investigate a single measure, the noise μ. The following concept will play an important role in this investigation.

3.15 Definition: An interval $(a, b]$ is left adjusted in an interval (c, d) iff $a = c$ and $b \leq d$. A sequence of disjoint intervals $\{(a_i, b_i]\}_{i=1}^n$ is left adjusted in an interval (c, d) iff $(a_1, b_1]$ is left adjusted in (c, d) and $(a_i, b_i]$ is left adjusted in (b_{i-1}, d) for $i = 2, 3, \ldots, n$.

3.16 Definition: Let μ be a measure defined on $(R, \mathcal{B})$. Let $0 \leq \lambda < 1$. An interval $(a, b]$ is (μ, λ)-minimal left adjusted in (c, d) iff $(a, b]$ is minimal left adjusted in (c, d), $\mu(a, b) \geq \lambda$, and if $\mu(c, b_1) \geq \lambda$ then $b_1 \geq b$. Let $\{\mu_i\}_{i=1}^n$ be

measures defined on $(R, \mathcal{B})$. A sequence $\{(a_i, b_i]\}_{i=1}^n$ is $(\{\mu_i\}_{i=1}^n, \lambda)$-minimal left adjusted in (c, d) iff $(a_1, b_1]$ is (μ_1, λ)-minimal left adjusted in (c, d) and (a_i, b_i) is (μ_i, λ)-minimal left adjusted in (b_{i-1}, d) for $i = 2, 3, \ldots, n$.

3.17 Lemma: Let $Y = (c, d)$. If μ is a totally finite measure on (Y, Y'), then for $0 < \lambda \leq \mu(Y)$ there exists an interval $(a, b]$ which is (μ, λ)-minimal left adjusted in Y.

Proof:

Let $A = \{b: \mu(c, b] \geq \lambda\}$. $A \neq \emptyset$ since $\lambda \leq \mu(Y)$. Let $b_o = \inf\{b: b \in A\}$. Let $\{b_n\}_{n=1}^{\infty}$ be a sequence contained in A which converges to b_o. It may be assumed with no loss of generality that $\{b_n\}_{n=1}^{\infty}$, is a decreasing sequence. Then $\{(c, b_n]\}_{n=1}^{\infty}$, is a decreasing sequence of intervals such that $\mu(c, b_n] < \infty$ for all n and $(c, b_o] = \bigcap_{i=1}^{\infty} (c, b_i]$. Therefore $\mu(c, b_o] = \lim_{n\to\infty} \mu(c, b_n] \geq \lambda$. Now it is clear by the nature of b_o that $(c, b_o]$ is (μ, λ)-minimal left adjusted.

3.18 Theorem: Let $S = \{X, \mu(.|x), (Y, Y')\}$ be a channel with Y connected. Let $0 \leq \lambda < 1$ be given. Let $(x_1, D_1), \ldots, (x_n, D_n)$ be a connected λ-code of length $N(S, \lambda, \mathcal{U})$. Then there exists a $(\{\mu(.|x_i)\}_{i=1}^n, 1-\lambda)$-minimal left adjusted sequence.

Proof:

It may be assumed that the sequence $\{D_i\}_{i=1}^n$ is ordered;

i.e. if one denotes $D_i = (a_i, b_i]$ then $i < j$ implies $a_i < a_j$. By lemma 3.17, there exists a $(\mu(.|x_1), 1-\lambda)$-minimal left adjusted interval in Y. Call it D_1'. Since D_1' is $(\mu(.|x_1), 1-\lambda)$-minimal left adjusted, it is clear $D_2 \subset Y-D_1'$; hence, $\mu(Y-D_1' \mid x_2) \geq 1-\lambda$. Thus, again by lemma 3.17, there exists and interval D_2' which is $(\mu(.|x_2), 1-\lambda)$-minimal left adjusted in $Y-D_1'$. It is clear that proceeding thusly one obtains the desired sequence $\{D_i'\}_{i=1}^{n}$.

3.19 Remark: In lemma 3.17 if $\mu << \gamma$ then $\mu(c, b_o] = \lambda$; hence in theorem 3.18, if $\mu(.|x) << \gamma$ for each $x \in X$ $\mu(D_i'|x_i) = 1-\lambda$ for $i = 1, 2, \ldots, n$.

3.20 Definition: Let $S = \{X, \mu(.|x), (Y, Y')\}$ be a channel with additive noise μ which is absolutely continuous with respect to γ. If Y is the real numbers, Y' the Borel sets, and X a closed interval contained in Y, then S is called a channel of type A.

Throughout the remainder of the Chapter the emphasis will be on channels of type A. In most of the analysis it will also be assumed that $f = d\mu/d\gamma$ is a bell function. Techniques for actually constructing connected λ-codes of length $N(S, \lambda, \mathcal{U})$ will now be presented. The first such construction is for an arbitrary channel of type A. Since $\mu(\{y\}) = 0$ the sets will be written as open intervals.

<u>3.21 Construction</u>: Let S be a channel of Type A. Let $0 \leq \lambda < 1$ be given. Let $X = [a, b]$. Let d_1 be such that $\mu(c, d_1)$ is $(\mu(.|a), 1-\lambda)$ minimal left adjusted. The remainder of the code is constructed inductively as follows.

Suppose members $\{x_i, D_i\}_{i=1}^{N}$ have been obtained. Let $D_N = (d_{N-1}, d_N)$. If $\mu((d_N, d)|b) < 1-\lambda$ the construction is completed; otherwise, d_{N+1} is obtained as follows. Let $\ell = \inf \{d^* - d_N |$ there exists $x \in [a, b]$ such that $\mu((d_N, d^*)|x) \geq 1-\lambda.\}$ Clearly $\mu((d_N, d_N + \ell - \varepsilon)|x) < 1-\lambda$ for all $\varepsilon > o$ and all $x \in X$. It will be shown that there exists $x \in X$ such that $\mu((d_N, d_N + \ell)|x) \geq 1-\lambda$. The n+1st element of the code will be chosen to be $\{x, (d_N, d_N + \ell)\}$. If there is more than one x such that $\mu((d_N, d_N + \ell)|x) \geq 1-\lambda$ then any such x may be chosen.

Let F(y) be the distribution function of μ. $\mu << \gamma$ implies F(y) is continuous. Observe that for any $x \in X$ $F_x(y) = F(y-x)$ where $F_x(y)$ is the distribution function of $\mu(.|x)$. Also if $x \in X$, $k > 0$ then $\mu((d_N, d_N + k)|x) = F(d_N - x + k) - F(d_N - x)$.

Let $\ell_n \searrow \ell$ then for each n there exists x_n such that $F(d_N - x_n + \ell_n) - F(d_N - x_n) \geq 1-\lambda$. $\{x_n\}_{n=1}^{\infty} \subset X$ hence is bounded and therefore has a limit point, say x'. Let

$(x_{n_k})_{n_k=1}^{\infty} \to x'$. Then $F(d_N - x' + \ell) - F(d_N - x')$

$= \lim_{n_k \to \infty} (F(d_N - x_{n_k} + \ell_{n_k}) - F(d_N - x_{n_k})) \geq 1-\lambda$. This proves the induction step. It is clear that the code constructed above has length $N(S, \lambda, \mathcal{U})$.

In the preceding construction the specific nature of the code is not readily apparent. In the case where f is a bell function one can construct the code so that it is more transparent, and, hence easier to manipulate.

3.22 Construction: Let S be a channel of type A. Let $f = d\mu/d\gamma$ be a bell function and let $0 \leq \lambda < 1$ be given. Let $X = [a, b]$ and $Y = (c, d)$ where both $c = -\infty$ and $d = +\infty$ are allowed.

Since f is a bell function, there exists, by theorem 3.11, a connected set (t_1, t_2) of minimal γ-measure such that $\mu(t_1, t_2) \geq 1-\lambda$. The numbers t_1 and t_2 will be used to construct a λ-code of length $N(S, \lambda, \mathcal{U})$.

Let (c, d_1) be the $(\mu(\cdot|a), 1-\lambda)$-minimal left adjusted interval in (c, d). This interval exists by lemma 3.17. The remainder of the λ-code is constructed as follows. Let the ith pair be denoted by $(x_i, (d_{i-1}, d_i))$. The i+1st pair, if it exists,

is constructed by

<u>Case 1</u>: If $d_i \leq a + t_1$ let (d_i, d_{i+1}) be the $(\mu(.|a), 1-\lambda)$-minimal left adjusted interval in (d_i, d) and let $x_{i+1} = a$.

<u>Case 2</u>: If $a + t_1 < d_i < b + t_1$ let d_{i+1} and $x_{i+1} = d_i - t_1$. By definition of t_1 and t_2, it is clear that (d_i, d_{i+1}) is $(\mu(.|x_{i+1}), 1-\lambda)$-minimal left adjusted in (d_i, d).

<u>Case 3</u>: If $d_i \geq b + t_1$ and $\mu(d_i, d)|b) < 1-\lambda$ the construction is completed; otherwise ($d_i \geq b + t_1$ and $\mu((d_i, d)|b) \geq 1-\lambda$) let (d_i, d_{i+1}) be the $(\mu(.|b), 1-\lambda)$-minimal left adjusted interval in (d_i, d) and let $x_{i+1} = b$.

The code constructed above will be labeled by $\{(x_i, D_i)\}_{i=1}^N$. It is clear that $\{D_i\}_{i=1}^N$ is $(\{\mu(.|x_i\}_{i=1}^N, 1-\lambda)$-minimal left adjusted in Y and that if $\{(\xi_i, B_i)\}_{i=1}^N$ is any set of pairs such that $\{B_i\}_{i=1}^N$ is $(\{\mu(.|\xi_i)\}_{i=1}^N, 1-\lambda)$-minimal left adjusted in Y then $b_N \geq d_N$ where $B_i = (a_i, b_i]$. Thus it follows from theorem 3.18 that $N = N(S, \lambda, \mathcal{U})$. Any member of the code constructed above of the form (a, D_i) will be called an <u>a-pair</u>; any member of the form (b, D_i) will be called a <u>b-pair</u>.

It is clear that the code constructed above is not, in general, the only connected code of length $N(S, \lambda, \mathcal{U})$. For reference later this code will be referred to as a <u>connected code of type 1</u>.

From the constructions outlined above it is clear that a more precise notation for the code is $\{x_i(\lambda), D_i(\lambda)\}_{i=1}^{N(S, \lambda, \mathcal{U})}$. However when there is no possibility of confusion the λ will be supressed.

3.23 Definition: Let $S = \{X, \mu(.|x), (Y, Y')\}$ be a channel of type A. Let $0 \leq \lambda < 1$ and let $\{x_i, D_i\}_{i=1}^N$ be a connected λ-code of type 1. This code will called full if $\mu((Y - \bigcup_{i=1}^{N} D_i)|x) = 0$ for $x \in X$. The term full code will always refer to a connected code of type 1 which is full.

It is easy to see that given any $N > N(S, 0, \lambda)$ there exists λ_N such that $N(S, \lambda_N, \mathcal{U}) = N$, and the λ_N-code is full iff given $\lambda < \lambda_N$ then $N(S, \lambda_N, \mathcal{U}) > N(S, \lambda, \mathcal{U})$.

3.24 Remark: Let $S = \{[a, b], \mu(.|x), (Y, Y')\}$ be a channel of type A with additive noise μ. Let $0 \leq \lambda < 1$ be given. Then there exists a channel S_1, more precisely an input alphabet $[a_1, b_1]$, such that the λ-code for S_1 is full. In order to verify this statement it suffices to demonstrate such an alphabet. To remove any possibility of ambiguity let the family of measures $\{\mu(.|x)|x \in R\}$ be defined by $\mu(A|x) = \mu(A-x)$ for all $x \in R$, $A \in Y'$. Since $\mu << \gamma$ there exists an interval $I = (t_1, t_2)$ such that $\mu(I) = 1-\lambda$ and given any interval J with $\gamma(J) < t_2 - t_1$

then $\mu(J) < 1-\lambda$. It is easy to see that there exists $y_1 > a$ such that $\mu(A|x) < 1-\lambda$ for all $x \geq y_1$, $A \subset (-\infty, a + t_2)$; similarly there exists $y_2 < b$ such that $\mu(A|x) < 1-\lambda$ for all $x \leq y_2$, $A \subset (b + t_1, \infty)$. Let $[a', b']$ be defined by $a' = a$, $b' = y_1 + (b - y_2) + 2(t_2 - t_1)$, and let S' be the channel with alphabet $[a', b']$. Let $\{x_i, D_i\}_{i=1}^N$ be the type 1 λ-code for S'. It is easy to see that if $D_i \subset (a + t_1, b + t_2)$ then $\gamma(D_i) = t_2 - t_1$. Let $\{x_i', D_i'\}_{i=1}^N$ be the code constructed analogously to $\{x_i, D_i\}_{i=1}^N$ where one uses right adjusted sequences. Let $\alpha = \sup \{y | y \in D_i, D_i \cap (-\infty, a + t_1) \neq \emptyset\}$; let $\beta = \inf \{y | y \in D_i', D_i' \cap (b + t_2, \infty) \neq \emptyset\}$. Given $\xi \geq b'$ let $\beta(\xi)$ be the point, considering the channel with input $[a, \xi]$, analogous to β. Then given $\xi_2 \geq \xi_1 \geq b'$ then $\beta(\xi_2) = \beta(\xi_1) + \xi_2 - \xi_1$. Thus one can clearly choose $b^* \geq b'$ such that $\beta(b^*) - \alpha$ is an integral multiple of $(t_2 - t_1)$. The type 1 λ-code for the channel with input alphabet $[a, b^*]$ is full.

In 3.21 it is clear that for a given $0 \leq \lambda < 1$ and $1 \leq i \leq N(S, \lambda, \mathcal{U})$ there may exist several $x \in X$ such that $\mu(D_i|x) \geq 1-\lambda$. We define $X_i^\lambda = \{x | x \in X \;\; \mu(D_i|x) \geq 1-\lambda\}$.

<u>3.25 Theorem</u>: Let S be a channel of type A. Suppose $N(S, \lambda, \mathcal{U}) = N(S, \lambda)$ for all $0 \leq \lambda < 1$. Let $0 \leq \lambda < 1$ be given such that the λ-code $\{x_i(\lambda), D_i(\lambda)\}_{i=1}^{N(S, \lambda, \mathcal{U})}$ is full. Then given $1 \leq i \neq j \leq N(S, \lambda, \mathcal{U})$ and $A_i \subset D_i(\lambda)$, $A_j \subset D_j(\lambda)$ such

that there exists $x_i^* \in X_i^\lambda$ for which $\mu(A_i|x_i^*) = \mu(A_j|x_i^*) \neq 0$ then $\mu(A_i|x_j^*) \leq \mu(A_j|x_j^*)$ for all $x_j^* \in X_j^\lambda$.

Proof:

Suppose there exists λ such that the λ-code is full and there exists $A_i \subset D_i(\lambda)$, $A_j \subset D_j(\lambda)$ for some $i \neq j$ such that there exists $x_i^* \in X_i^\lambda$, $x_j^* \in X_j^\lambda$ such that $\mu(A_i|x_i^*) = \mu(A_j|x_i^*) > 0$ and $\mu(A_i|x_j^*) > \mu(A_j|x_j^*)$. Then $\mu((D_i(\lambda) - A_i) \cup A_j)|x_i^*) = 1-\lambda$ and $\mu((D_j(\lambda) - A_j) \cup A_i)|x_j^*) > 1-\lambda$. Hence choosing $D_i'(\lambda) = D_i(\lambda) - A_i) \cup A_j$ and $D_j'(\lambda) \subset (D_j(\lambda) \cup A_i) - A_j$ such that $\mu(D_j'|x_j^*) = 1-\lambda$ one can obtain a λ-code $\{\xi_i, B_i\}_{i=1}^{N(S, \lambda, U)}$ such that there exists at least one $x \in X$ such that $\mu(Y - \bigcup_{i=1}^{N(S, \lambda, U)} B_i|x) > 0$. It follows by the absolute continuity of λ, (More precisely by the fact that the end-point of each B_i is a continuous function of λ) that there exists $\lambda_o < \lambda$ such that $N(S, \lambda_o) > N(S, \lambda_o, U)$. But this is a a contradiction of the hypothesis that $N(S, \lambda) = N(S, \lambda, U)$ for all $0 \leq \lambda < 1$.

3.26 Corollary: If there exists a full λ-code $\{x_i, D_i\}_{i=1}^{N(S, \lambda, U)}$ and for some $1 \leq i \neq j \leq N(S, \lambda, U)$ there is

$A_i \subset D_i$, $A_j \subset D_j$ such that $\mu(A_i|x_i^*) \neq 0 \neq \mu(A_j|x_i^*)$ and $f(y_i|x_j^*)/f(y_i|x_i^*) > f(y_j|x_j^*)/f(y_j|x_i^*)$ for almost all $[\gamma]$ $y_i \epsilon A_i$, $y_j \epsilon A_j$ and some $x_i^* \epsilon X_i^\lambda$, $x_j^* \epsilon X_j^\lambda$ then there exists $0 < \lambda_1 < 1$ such that $N(S, \lambda_1) > N(S, \lambda_1, U)$.

Proof:

Let $B_i \subset A_i$, $B_j \subset A_j$ such that $\mu(B_i|x_i^*) = \mu(B_j|x_i^*) > 0$. Then

$$\mu(B_i|x_j^*) = \int_{B_i} f(y|x_j^*)\,d\gamma = \int_{B_i} f(y|x_i^*)\frac{f(y|x_j^*)}{f(y|x_i^*)}\,d\gamma >$$

$$\int_{B_j} f(y|x_i^*)\frac{f(y|x_j^*)}{f(y|x_i^*)}\,d\gamma = \mu(B_j|x_j^*)$$

3.27 Corollary: Suppose $f = d\mu/d\gamma$ is a bell function and is unbounded. Then there exists a channel S of type A with additive noise μ and $0 < \lambda < 1$ such that $N(S, \lambda) > N(S, \lambda, U)$.

Proof:

It follows from the definition of a bell function that there exists at most one point, y_0, such that $\overline{\lim_{y \to y_0}} f(y) = \infty$. One may as well assume that $y_0 = 0$.

Observe that it suffices to show that there exists a channel $S = \{[a, b], \mu(\cdot|x), (Y, Y')\}$ and a full λ-code $\{x_i, D_i\}_{i=1}^{N}$ such that either (1) $b_1 < a < x_2$ or (2) $x_{N-1} < b < a_N$ where $D_i = (a_i, b_i)$. In the first case $f(y|a)/f(y|x_2)$ is unbounded on D_2 while this function is bounded on a subset of D_1 of positive $\mu(\cdot|x_2)$ measure. Similarly, in the second case, $f(y|b)/f(y|x_{N-1})$ is unbounded on D_{N-1} while this function is bounded on a subset of D_N of positive $\mu(\cdot|x_{N-1})$ measure.

If either $\mu(-\infty, 0) = 0$ or $\mu(0, \infty) = 0$ then, for a given channel S, more precisely for a given input alphabet $[a, b]$, it is clear that one can choose $0 < \lambda < 1/2$ such that the λ-code is full. Thus either (1) or (2) must hold.

Suppose $(-\infty, 0) \neq 0 \neq \mu(0, \infty)$. One may assume with no loss of generality that $\mu(-\infty, 0) \geq \mu(0, \infty)$. Observe that as $1-\lambda$ increases $t_1(\lambda)$ decreases. Thus there exists $1-\lambda < \mu(-\infty, 0)$ such that $t_1(\lambda) < a_1$. By 3.24 there exists a channel S such that the λ-code is full. Since $t_1(\lambda) < a_1 < a$ (1) must hold.

<u>3.28 Remark</u>: Sufficient conditions for $N(S, \lambda) = N(S, \lambda, U)$ will now be shown. Throughout the remainder of this chapter it will be assumed that f is bounded and $f(0) \geq f(y)$ for all $y \in Y$.

The importance of the behavior of the ratio $f(y)/f(y-d)$ was partially shown in 3.26 and 3.27 in the form of necessary

conditions. The following theorem, which continues this investigation, will be used to derive a sufficient condition for $N(S, \lambda) = N(S, \lambda, \mathcal{U})$.

<u>3.29 Definition</u>: Given sets B_1 and B_2 then B_1 is <u>less than</u> B_2, $B_1 < B_2$, if given any $y_1 \varepsilon B_1$, $y_2 \varepsilon B_2$ then $y_1 < y_2$.

<u>3.30 Theorem</u>: Let $x_1, x_2 \varepsilon X$ with $x_2 = x_1 + d$, $d > 0$. Let $0 < \lambda < 1$. If $f(y)/f(y-d)$ is a decreasing function of y then given any A_1, A_2 with $A_1 \cap A_2 = \emptyset$ and $\mu(A_i|x_i) \geq 1-\lambda$, $i = 1, 2$ there exists $B_1, B_2 \subset A_1 \cup A_2$ with $B_1 < B_2$ and $\mu(B_i|x_i) \geq 1-\lambda$, $i = 1, 2$.

<u>Proof</u>:

Observe that $f(y)/f(y-d) = d\mu(.|x_1)/d\mu(.|x_2)$. Let $\xi = \inf\{y|\mu((A_1 \cup A_2) \cap (-\infty, y)|x_1) \geq 1-\lambda\}$. Let $B_1 = (A_1 \cup A_2) \cap (-\infty, \xi)$. Then, since $\mu << \gamma$, $\mu(B_1|x_1) = 1-\lambda$. By 3.5 B_1 is a set of minimal $\mu(.|x_2)$ measure such that $\mu(B_1|x_1) \geq 1-\lambda$. Let $B_2 = (\xi, \infty) \cap (A_1 \cup A_2)$. $\mu(B_1|x_2) + \mu(B_2|x_2) = \mu(A_1|x_2) + \mu(A_2|x_2)$. Hence, since $\mu(B_1|x_2) \leq \mu(A_1|x_2)$, $\mu(B_2|x_2) \geq \mu(A_2|x_2) \geq 1-\lambda$.

<u>3.31 Lemma</u>: Let $0 \leq \lambda < 1$. Suppose that given any $x_1, x_2 \varepsilon X$ with $x_1 < x_2$ such that there exists A_1, A_2 with $A_1 \cap A_2 = \emptyset$ and $\mu(A_i|x_i) \geq 1-\lambda$, $i=1, 2$ then there exists $B_1, B_2 \subset A_1 \cup A_2$ with $B_1 < B_2$, $B_1 \cap B_2 = \emptyset$, and $\mu(B_i|x_i) \geq 1-\lambda$, $i=1, 2$. Then $N(S, \lambda) = N(S, \lambda, \mathcal{U})$.

Proof:

Let $\{x_i, A_i\}_{i=1}^N$ be a λ-code of length $N(S, \lambda)$. It will be assumed that the indexing is such that $i < j$ implies $x_i < x_j$. By hypothesis there exists a code $\{x_i, A_i'\}_{i=1}^N$ with $A_1' \cup A_N' \subset A_1 \cup A_N$ and $A_1' < A_N'$. By the same argument there is a code $\{x_i, A_i''\}_{i=1}^N$ with $A_1'' < A_{N-1}''$ and $A_1'' < A_N''$. It is easy to see that repetition of this logic proves the existance of a code $\{x_i, B_i\}_{i=1}^N$ with $B_1 < B_2 < B_i$ for $3 \leq i \leq N$. It is now clear that there exists a λ-code $\{x_i, C_i\}_{i=1}^N$ with $C_i < C_j$ whenever $i < j$.

Let $a_i = \inf \{y|y \in C_i\}$; $b_i = \sup \{y|y \in C_i\}$. Let $D_i = (a_i, b_i)$. Then $\mu(D_i|x_i) \geq \mu(C_i|x_i) \geq 1-\lambda$ and $D_i \cap D_j = \emptyset$ if $i \neq j$. Thus, $\{x_i, D_i\}_{i=1}^N$ is a connected λ-code of length $N(S, \lambda)$.

The following theorem is an immediate consequence of the two lemmas.

3.32 Theorem: Let S be a channel of type A with $f = d\mu/d\gamma$ a bell function. If f is log-concave then $N(S, \lambda) = N(S, \lambda, \mathcal{U})$ for all $0 \leq \lambda < 1$. In particular, $N(S, \lambda) = N(S, \lambda, \mathcal{U})$ for all $0 \leq \lambda < 1$ if S has additive Gaussian noise.

It will now be shown that there exists channels of type A with $f = d\mu/d\gamma$ a bell function for which there exists $0 < \lambda < 1$ such that $N(S, \lambda) > N(S, \lambda, \mathcal{U})$.

<u>3.33 Definition</u>: Let F be the set of all functions of the following kind. Given $0 < s < 1$, $1 < t < 1/s$, define

$$f(y) = t \qquad \text{if } y \in (-s/2, s/2)$$

$$f(y) = u = \frac{1-st}{1-s} \text{ if } y \in (-1/2, -s/2) \cup (s/2, 1/2)$$

$$= 0 \qquad \text{elsewhere.}$$

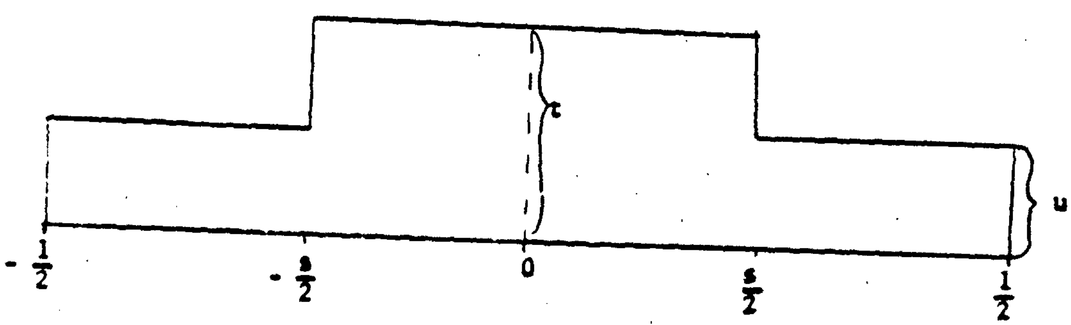

Figure 1: A Typical Member of F

<u>3.34 Remark</u>: Given $f \in F$ then f induces a probability measure, μ_f, on the Borel sets in a natural way, i.e.

$$\mu_f(A) = \int_A f(y)\,d\gamma$$

for all $A \in \mathcal{B}$.

<u>3.35 Definition</u>: Given $\ell > 0$ and $f \in F$ let $S(f, \ell)$ denote

any channel $\{[a, b], \mu(.|x), (Y, Y')\}$ with additive noise μ_f and $b-a = \ell$.

It will now be shown that for certain members of the family $\{S(f, \ell) | f \in F, \ell > 0\}$ $N(S, \lambda) = N(S, \lambda, U)$ for all $0 \leq \lambda < 1$ while for other members of the family equality does not always hold.

<u>3.36 Remark</u>: Let S be a channel of type A. Given $y \in Y$ define $f^*(y) = \{\sup f(y|x) | x \in X\}$. Let $F(y) = \int_Y f^*(y) d\gamma$. Then given $0 \leq \lambda < 1$ and any λ-code $\{x_i, A_i\}_{i=1}^N$ then

$$\sum_{i=1} \mu(a_i | x_i) \leq F(Y).$$

[illegible] <u>Lemma</u>: Given $\ell > 0$, $f \in F$, $N(S(f, \ell), \lambda) = N(S(f, \ell), \lambda, U)$ f[illegible] $\lambda \geq 1-st$.

<u>Proof</u>:

Let $\lambda_1 \geq 1-st$ such that λ_1-code is full. Let $\{x_i, D_i\}_{i=1}^N$ be a full code. Then $\sum_{i=1}^N \mu(D_i | x_i) = F(Y)$. Hence $N(S(f, \ell), \lambda_1) = N(S(f, \ell), \lambda_1, U)$. Suppose there exists $\lambda_2 < \lambda_1$ such that $N(S(f, \ell), \lambda_2) = N(S(f, \ell), \lambda_1) = N$. Let

$\{\xi_i, B_i\}_{i=1}^{N}$ be a λ_2-code. Then $\sum_{i=1}^{N} \mu(B_i|\xi_i) \geq N(1-\lambda_2) > N(1-\lambda_1)$ $= F(Y)$. Which is a contradiction. Hence $\lambda < \lambda_1$ implies $N(S(f, \ell), \lambda) < N(S(f, \ell), \lambda_1)$. This completes the proof.

<u>3.38 Remark</u>: Let $\{x_i, A_i\}_{i=1}^{N}$ be a full λ-code with $\lambda < 1\text{-}st$. Let k be the number of a-pairs. Let $A_i = (a_i, b_i)$ Observe that $b_i - a_i > s$. Suppose $x_{k+1} \neq b$. Then X_{k+1}^{λ} is an interval, say $[\alpha_1, \alpha_2]$. In fact, see Figure 2, $\alpha_1 = b_2 + s/2$ and $\alpha_2 - \alpha_1 = (1-\lambda-st)/u$. If $b_k - (a - s/2) < \frac{1-s}{2}$ there exists an interval $(\beta, a-s/2)$ such that $f(y|\alpha_1) = u$ for all $y \in (\beta, a-s/2)$, see Figure 3. Thus under these conditions it is easy to see, Figure 4, that one can choose $x'_{k+1} \in X_{k+1}^{\lambda}$ such that there exists $y_1 \in A_k$, $y_2 \in A_{k+1}$ such that $f(y_1|a) = u = f(y_1|x'_{k+1})$ while $f(y_2|a) = t$ and $f(y_2|x_{k+1}) = u$. It follows from 3.25 that, under the conditions specified above, there exists $0 \leq \lambda < 1$ such that $N(S(f, \ell), \lambda) > N(S(f, \ell), \lambda, \mathcal{U})$.

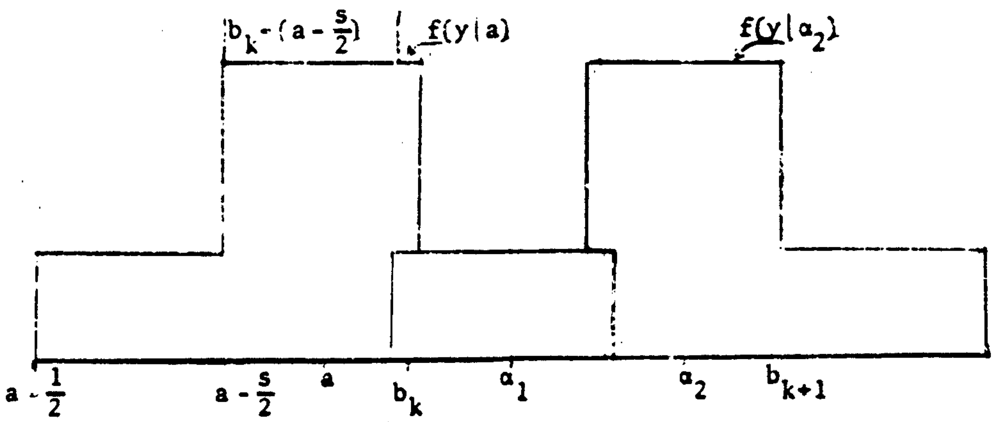

Figure 2. Definition of the Interval (α_1, α_2)

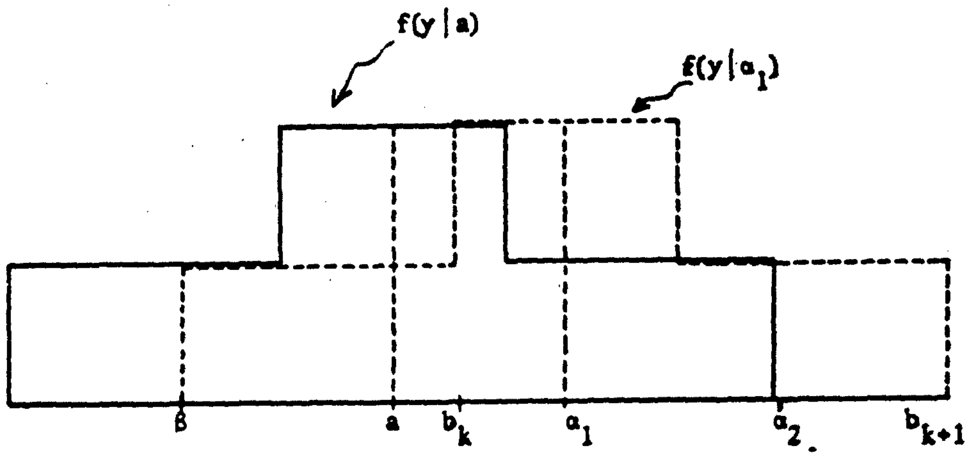

Figure 3: Definition of the Interval (β, $a-s/2$)

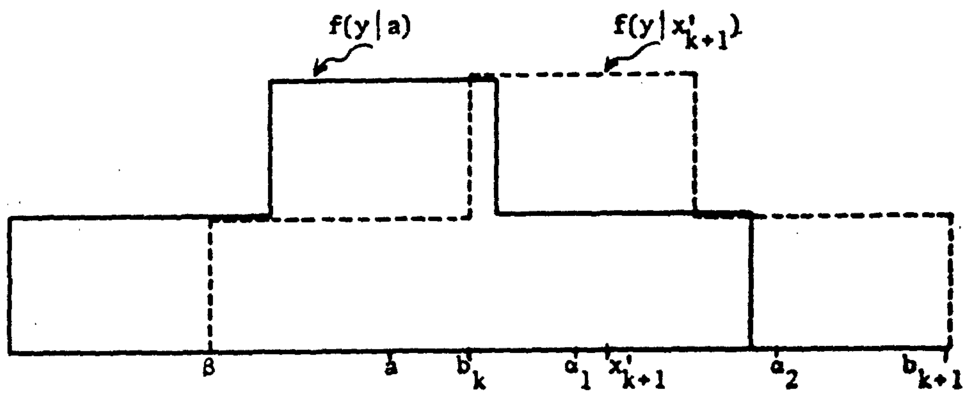

Figure 4: A Choice for $f(y|x'_{k+1})$

<u>3.39 Lemma</u>: If $st < 1/3$ there exists $0 < \lambda < 1$ such that $N(S(f, \mathcal{L}), \lambda) > N(S(f, \mathcal{L}), \lambda, \mathcal{U})$.

<u>Proof:</u>

If $b-a < s$ let λ be such that $N(S(f, \mathcal{L}), \lambda, \mathcal{U}) = 3$ and the λ-code $\{x_i, A_i\}_{i=1}^N$ is full. Let $A_i = (a_i, b_i)$. Observe that $b_1 < a + s/2$. One can assume that $x_2 = (a + b)/2$. Now $b_1 - (a - s/2) < s < \frac{1-s}{2}$. Thus, for this case, the conclusion follows from 3.38.

Suppose $b - a \geq s$. Let $\lambda_o = \max\{\lambda \mid \lambda < 1-st$ and the λ-code is full$\}$. Let $\{x_i, A_i\}_{i=1}^N$ be a full λ_o-code where $A_i = (a_i, b_i)$ and let k be the number of a-pairs. Then either (1) $b_k < a + s/2$, or (2) $b_k > a + s/2$, or (3) $b_k = a + s/2$. Observe $b_k \geq a + s/2$ implies $k \geq 2$ since $b_k \geq a + s/2$, $k = 1$, and $s < 1/3$ implies $b_2 \geq a - s/2 + s + \frac{1-s}{2} + s > a - s/2 + 3s$; but, this indicates that $b_2 - (a - s/2) > 3s$ which says that there exists $\lambda_o < \lambda_1 < 1 - st$ such that $N(S, \lambda_1, \mathcal{U}) \geq N(S, \lambda_o, \mathcal{U}) + 1$ which contradicts the definition of λ_o.

If (1) is true then $b_k - (a - s/2) < s < \frac{1-s}{2}$ and the conclusion follows from 3.38. If (2) is true then there exists $\xi \in (a, b)$ such that $\xi \in X_k^\lambda$. Now, replacing (x_k, A_k) by (ξ, A_k),

one obtains a code with k-1 a-pairs and $b_{k-1} \leq a - s/2$. Again the conclusion follows by 3.38. If (3) is true let $\lambda_1 = \max\{\lambda \mid \lambda < \lambda_o$ and the λ-code is full$\}$. Let $\{\xi_i, B_i\}_{i=1}^M$ be a full λ_1-code and let n be the number of a-pairs. Let $A_i' = (a_i', b_i')$. Then either (1') $b_n' < a + s/2$ or (2') $b_n' > a + s/2$. The conclusion follows by the same arguments used above.

3.40 Theorem: $N(S(t, \mathcal{L}), \lambda) = N(S(f, \mathcal{L}), \lambda, \mathcal{U})$ for all $0 \leq \lambda < 1$ iff given $\lambda < 1 - st$ such that the λ-code is full then either $N(S(f, \mathcal{L}), \lambda, \mathcal{U}) = 2$ or $\lambda \leq 1 + st - \frac{t}{2} - \frac{1}{2}$.

Proof:

The condition $\lambda \leq 1 + st - \frac{t}{2} - \frac{1}{2}$ is equivalent to $1 - \lambda \geq \frac{t}{2} + \frac{1}{2} - st = \frac{1-st}{2} + (\frac{1-s}{2})t$.

To show that the condition is necessary assume $N(S(f, \mathcal{L}), \lambda, \mathcal{U}) = N(S(f, \mathcal{L}), \lambda)$ for all $0 \leq \lambda < 1$. By 3.39 $st \geq 1/3$. Let $\lambda_o = \max\{\lambda \mid \lambda < 1 - st$ and the λ-code is full$\}$. Let $\{x_i, D_i\}_{i=1}^N$ be a full λ_o-code and let $D_i = (a_i, b_i)$. Suppose $N(S(f, \mathcal{L}), \lambda_o, \mathcal{U}) > 2$. Since $1 - \lambda_o > st \geq 1/3 \geq \mu((a - \frac{1}{2}, a - s/2) \mid a)$, $b_1 > a - s/2$. If $b_1 < a + s/2$ then $1 - \lambda_o = \mu((a - \frac{1}{2}, b_1) \mid a) = \frac{1-st}{2} + (b_1 - (a - \frac{s}{2}))t$ which, by 3.38 is greater than or equal to $\frac{1-st}{2} + (\frac{1-s}{2})t$. If $b_1 \geq a+s/2$

it is easy to see that $b_i - a_i \geq \frac{1-s}{2} + s$ for $1 \leq i \leq N$. Thus $b_2 \geq a + s/2 + \frac{1-s}{2} + s$; hence $b_2 - (a - s/2) \geq \frac{1-s}{2} + 2s$. It follows by the maximality of λ_0 that $\frac{1-s}{2} < s$. Hence,
$1 - \lambda = \mu(D_1|a) \geq \mu((a - \frac{1}{2}, a + s/2)|a) = \frac{1-st}{2} + st > \frac{1-st}{2} + (\frac{1-s}{2})t$. Thus the condition is necessary.

To show that the condition is sufficient let $0 < \lambda \leq 1 + st - \frac{t}{2} - \frac{1}{2}$. Let $x_1 < x_2$ and let $A_1, A_2 \varepsilon Y'$, $A_1 \cap A_2 = \emptyset$. If $\frac{1-s}{2} > s$ and $x_2 - x_1 < \frac{1-s}{2}$ then $\mu(A_1|x_1) + \mu(A_2|x_2) < 1 - st + (\frac{1-s}{2} + s)t < 2(1-\lambda)$. Thus if $\mu(A_i|x_i) \geq 1-\lambda$, $i = 1, 2$ then either (1) $x_2 - x_1 \geq \frac{1-s}{2}$ or (2) $x_2 - x_1 < \frac{1-s}{2} < s$.

<u>Case 1</u>: $x_2 - x_1 < \frac{1-s}{2} < s$. It is easy to see that $x_1 + s/2 > x_2 - s/2$. Let $f^*(y) = \max (f(y|x_1), f(y|x_2))$ and let $\mu^*(A) = \int_A f^*(y)dy$ for all $a \varepsilon Y'$. Observe that given $x_2 - s/2 < y < x_1 + s/2$ then $\mu^*(A_1 \cup A_2) = \mu((A_1 \cup A_2) \cap (-\infty, y)|x_1) + \mu((A_1 \cup A_2) \cap (y, \infty)|x_2)$. Let $\xi = \inf \{y|\mu((A_1 \cup A_2) \cap (-\infty, y)|x_1) \geq 1 - \lambda$. Let $B_1 = (A_1 \cup A_2) \cap (-\infty, \xi)$. Then $\mu(B_1|x_1) = 1-\lambda$. Let $B_2 = (A_1 \cup A_2) \cap (\xi, \infty)$. Observe that $\mu^*((A_1 \cup A_2) \cap (-\infty, x_2 - s/2)) < 1-\lambda$, and $\mu^*((A_1 \cup A_2) \cap (x_1 + s/2)) < 1-\lambda$. If follows, since $\mu^*(A_1 \cup A_2) \geq 2(1-\lambda)$, that $x_2 - s/2 < \xi < x_1 + s_2$. Hence, $2(1-\lambda) \leq \mu^*(A_1 \cup A_2) = \mu(B_1|x_1) + \mu(B_2|x_2)$. Therefore, $\mu(B_2|x_2) \geq 1-\lambda$.

<u>Case 2</u>: $x_2 - x_1 \geq \frac{1-s}{2}$. Then $f(y|x_1)/f(y|x_2)$ is a decreasing function of y. Thus, by 3.30, there exists $B_1 < B_2 \subset A_1 \cup A_2$ such that $\mu(B_i|x_i) \geq 1-\lambda$, i = 1, 2.

If follows from 3.31 that $N(S(f, \mathcal{L}), \lambda, \mathcal{U}) = N(S(f, \mathcal{L}), \lambda)$ for all $\lambda \leq 1 + st - \frac{t}{2} - \frac{1}{2}$.

Suppose $N(S(f, \mathcal{L}), \lambda, \mathcal{U}) = 2$ and the λ-code $\{x_i, A_i\}_{i=1}^{2}$ is full then $x_1 = a$, $x_2 = b$ and $A_1 = (a - \frac{1}{2}, \frac{b+a}{2}]$. If $b - a \leq s$ then $\mu(A_1|a) + \mu(A_2|b) = 1 + (b-a)t = F(y)$. If follows from 3.36 that $N(S(f, \mathcal{L}), \lambda) = N(S(f, \mathcal{L}), \lambda, \mathcal{U})$. If $b - a > s$ then $\frac{b+a}{2} > a + s/2$. Hence if A is any set of minimal γ-measure such that $\mu(A) \geq 1-\lambda$ then $\mu(A) = (1-\lambda-st)/u = \frac{1}{2}(b+a+1)$. Hence $N(S(f, \mathcal{L}), \lambda) = 2$.

By lemma 3.37 $N(S(f, \mathcal{L}), \lambda) = N(S(f, \mathcal{L}), \lambda, \mathcal{U})$ for all $\lambda \geq 1-st$. This completes the proof.

<u>3.41 Corollary</u>: Given $\mathcal{L} > 0$ there exists $f \varepsilon F$ such that $N(S(f, \mathcal{L}), \lambda) > N(S(f, \mathcal{L}), \lambda, \mathcal{U})$ for some $0 < \lambda < 1$.

<u>Proof</u>:

Choose $f \varepsilon F$ such that $st < 1/3$.

<u>3.42 Corollary</u>: Given $f \varepsilon F$ with $st > 1/3$ there exists $\mathcal{L} > 0$ such that $N(S(f, \mathcal{L}), \lambda) = N(S(f, \mathcal{L}), \lambda, \mathcal{U})$ for all $0 < \lambda < 1$.

Proof:

Choose $\ell = 3s - \frac{1}{t}$. Let $\alpha = a + s - \frac{1}{2t}$, $\beta = b - s + \frac{1}{2t}$. Then $\mu((a - \frac{1}{2}, \alpha)|a) = \mu((\beta, b + \frac{1}{2})|b) = st.$. $\beta - \alpha = s$. Hence, given $\lambda < 1 - st$ such that the λ-code is full then $N(S(f, \ell), \lambda, U) = 2$.

3.43 Corollary: Given $f \varepsilon F$ with $1/3 < s < \frac{t+1}{4t}$ there exists $\ell_1 > 0$ such that $N(S(f, \ell_1), \lambda) > N(S(f, \ell_1), \lambda, U)$ for some $0 < \lambda < 1$ and there exists $\ell_2 > 0$ such that $N(S(f, \ell_2), \lambda) = N(S(f, \ell_2), \lambda, U)$ for all $0 < \lambda < 1$.

Proof:

The second conclusion is an immediate consequence of Corollary 3.42. To show the first conclusion observe that $s < \frac{t+1}{4t}$ implies that $1 - st > 1 + st - \frac{t}{2} - \frac{1}{2}$. Now, given $1 - st > \lambda > 1 + st - \frac{t}{2} - \frac{1}{2}$ one can choose $\ell_1 > 0$ such that the λ-code is full and $N(S(f, \ell_1), \lambda, U) > 2$.

3.44 Corollary: $N(S(f, \ell), \lambda) = N(S(f, \ell), \lambda, U)$ for all $\ell > 0$ and all $0 < \lambda < 1$ iff $s \geq \frac{t+1}{4t}$.

Proof:

The necessity of the condition is proved in 3.43. Suppose $s \geq \frac{t+1}{4t}$. Then $1 - st \leq 1 + st - \frac{1}{2} - \frac{t}{2}$.

3.45 Conjecture: It has been shown that if f is a bell function one may still have $N(S, \lambda) > N(S, \lambda, U)$ for some $0 \leq \lambda < 1$. The author has been unable to formulate a proof that if $N(S, \lambda) = N(S, \lambda, U)$ for all $0 \leq \lambda < 1$ then f must be a bell function. However, with the aid of theorem 3.25 many non-bell functions have been investigated and in all cases it has been possible to find a λ such that $N(S, \lambda) > N(S, \lambda, U)$. This, along with intuitive feeling, has led to the conjecture that $N(S, \lambda) = N(S, \lambda, U)$ for all $0 \leq \lambda < 1$ implies f is a bell function. Examples are listed below of channels of type A where f is not a bell function. It will be observed that in each example a family of functions (hence, a family of channels) is defined, and, in each example, given $\varepsilon > o$ there exists a member f of the family and a bell function g such that $|f(y) - g(y)| < \varepsilon$ for all $y \varepsilon Y$. Moreover, $N(S(g, \ell), \lambda) = N(S(g, \ell), \lambda, U)$ for all $0 \leq \lambda < 1$.

3.46 Example: Given $0 < \beta < 1$, $0 < \delta < 1/2$. Let

$$
\begin{aligned}
f(y) &= 1-\beta && \text{for } y\ \varepsilon(-\ \delta/2,\ \delta/2) \\
&= 1 + \frac{\delta\beta}{1-\delta} && \text{for } y\ \varepsilon(-\ 1/2,\ -\ \delta/2) \cup (\delta/2,\ 1/2) \\
&= 0 && \text{otherwise}
\end{aligned}
$$

Let $X = [0, 1/2]$, let $\lambda = 1 - \mu_f(-\ 1/2,\ 1/4)$. Then $N(S, \lambda, U) = 2$ and the λ-code is full. Observe that $f(y|1/2) > f(y|o)$ for all $y\ \varepsilon(0, \delta/2)$ and $f(y|o) > f(y|1/2)$ for all

$y \in (1/2 - \delta/2,\ 1/2)$. Hence, by 3.25 there exists $0 \le \lambda < 1$ such that $N(S, \lambda) > N(S, \lambda, U)$.

Define $g(y) = 1$ for $y \in (-1/2,\ 1/2)$

$= 0$ otherwise

Then, for $0 < \beta < \varepsilon$, $|f(y) - g(y)| < \varepsilon$. It is clear that $N(S(g, \ell), \lambda) = N(S(g, \ell), \lambda, U)$ for all $0 \le \lambda < 1$.

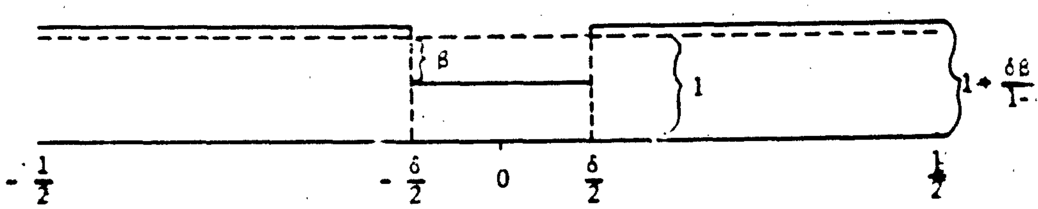

Figure 5. The First Counter Example.

3.47 Example: Given $f \in F$ with $s \ge (\frac{t+1}{4t})$. Define M_f to be the family of all function defined by: Given $0 < m < u(\frac{1+s}{2})$

define $g(y) = -my + u - m(\frac{1+s}{4})$ for $y \in (-1/2,\ -s/2)$

$= my + u - m(\frac{1+s}{4})$ for $y \in (s/2,\ 1/2)$.

$= f(y)$. otherwise

Given $\lambda < 1 - \frac{st}{2}$ such that there exists a full λ-code $\{x_i, D_i\}_{i=1}^N$ then there exists $1 < j < N$ such that $a - s/2 < y < a + s/2$ for all $y \varepsilon D_j$. Hence there exists $x_j^* \varepsilon X_j^\lambda - \{a\}$ such that $|a - x_j^*| < \frac{1-s}{2}$. Hence there exists $y_1 \varepsilon D_1$ such that $f(y_1|x_j^*) > f(y_1|x_1)$. Now $f(y|x_1) = f(y|x_j)$ for all $y \varepsilon D_j$. Thus by 3.25 $N(S, \lambda) > N(S, \lambda, \mathcal{U})$ for some $0 < \lambda < 1$.

Given $\varepsilon > 0$, then for $0 < m < \varepsilon$, $|f(y) - g(y)| < \varepsilon$ for all $y \varepsilon Y$. $N(S(f, \mathcal{L}), \lambda) = N(S(f, \mathcal{L}), \lambda, \mathcal{U})$ for all $0 < \lambda < 1$ by 3.44.

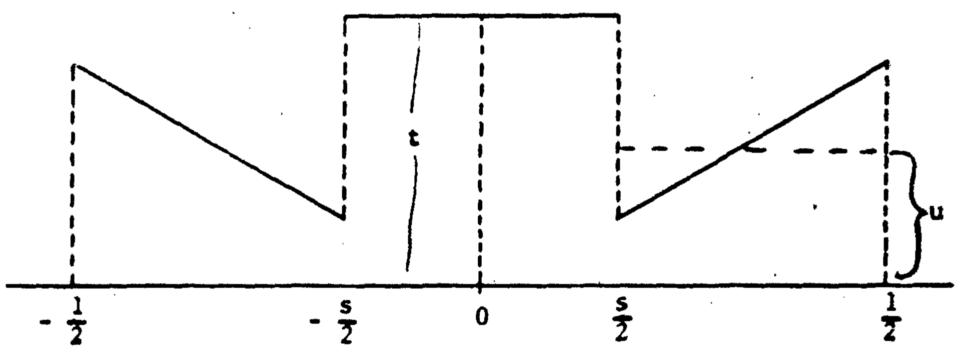

Figure 6. The Second Counter Example.

CHAPTER IV

ADMISSIBLE MEASURE

It is reasonably clear that in formulating necessary and sufficient conditions for $N(S, \lambda) = N(S, \lambda, \mathcal{U})$ one need not worry about every measure in the family $M = \{\mu(.|x): x \in X\}$. In fact for a fixed value of λ, say λ_o, if it is known that there exists a connected λ_o-code, $\{(x_i, D_i)\}_{i=1}^N$, of length $N(S, \lambda_o)$, then $N(S, \lambda_o, \mathcal{U}) = N(S, \lambda_o)$ regardless of the nature of the measures in the set $\{\mu(.|x): x \in (X - \{x_i\}_{i=1}^N)\}$. In this chapter a method will be defined for delineating those measures which cannot affect the value of $N(S, \lambda, \mathcal{U})$

Throughout this chapter both the input and the output will be subsets of the reals; γ will be Lebesgue measure; and $\mathcal{U}$ will be the usual topology for the reals.

<u>4.1 Definition</u>: A measure $\mu(.|x)$ is <u>admissible</u> if and only if there exists a connected set $A \in Y'$ such that $\mu(A|x) > 0$, $\mu(A|x') \leq \mu(A|x)$ for all $x' \in X$, and there exists $x'' \in X$ such that $\mu(A^c|x'') > 0$.

<u>4.2 Definition</u>: Given $0 \leq \lambda < 1$ a measure $\mu(.|x)$ is λ-admissible if and only if there exists a connected set $A \in Y'$

such that $\mu(A|x) \geq 1-\lambda$, $\mu(A|x') \leq \mu(A|x)$ for all $x' \epsilon X$, and there exists $x'' \epsilon X$ such that $\mu(A^c|x'') > 0$.

<u>4.3 Theorem</u>: Let $X^* = \{x \epsilon X: \mu(.|x)$ is admissible$\}$. Let $X_\lambda = \{x \epsilon X: \mu(.|x)$ is λ-admissible$\}$. Let $\{\lambda_n\}_{n+1}^{\infty}$ be a sequence such that $0 \leq \lambda_n < 1$ which converges to one. Then $X^* = \bigcup_{i=1}^{\infty} X_{\lambda_n}$

<u>Proof</u>:

Any λ-admissible measure is clearly admissible. Hence $\bigcup_{n=1}^{\infty} X_{\lambda_n} \subset X^*$. Let $x \epsilon X^*$, then there exists a connected set $A \epsilon Y'$ such that $\mu(A|x') \leq \mu(A|x)$ for all $x' \epsilon X$, $\mu(A^c|x'') > 0$ for some $x'' \epsilon X$, and $\mu(A|x) > 0$. Thus, there exists n such that $\mu(A|x) > 1- \lambda_n$ hence $x \epsilon X_{\lambda_n}$.

<u>4.4 Remark</u>: If the channel is semi-continuous, then for any $0 \leq \lambda < 1$ and any connected λ-code $(x_1, D_1), \ldots, (x_n, D_n)$, there exists connected a λ-code $(x_1', D_1), \ldots, (x_n', D_n)$ where $x_i' \epsilon X_\lambda$ for $i = 1, 2, \ldots, n$ unless X_λ is empty which implies $N(S, \lambda, \mathcal{U}) = 1$. However, for the general channel this need not hold since maximums may not be attained. The following theorem and corollary show that this is true for certain interesting special cases.

4.5 Theorem: Let $S = \{X, \mu(.|x), (Y, Y')\}$ be a given channel; let $S^* = \{X^*, \mu(.|x), (Y, Y')\}$. Then if X is compact and $\mu(.|x)$ is continuous in x, i.e. $\mu(A|x)$ is a continuous function of x for each $A\varepsilon Y'$, then $N(S^*, \lambda, U) = N(S, \lambda, U)$ whenever $N(S, \ , U) > 1$.

Proof:

Suppose $N(S, \lambda, U) \geq 2$. Let $\{x_i, D_i\}_{i=1}^{N(S, \lambda, U)}$ be a connected λ-code of length $N(S, \lambda, U)$. Let $1 \leq i \leq N(S, \lambda, U)$. Since $\mu(D_i|x)$ is a continuous function of x and X is compact $\mu(D_i|x)$ has a maximum at, say, $x'\varepsilon X$. $D_j \subset D_i^c$ for all $j \neq i$; hence, there exists $x''\varepsilon X$ such that $\mu(D_j^c|x'') > 0$. Therefore, $x_i'\varepsilon X$. If follows that $N(S^*, \lambda, U) = N(S, \lambda, U)$.

4.6 Corollary: Let S be the channel definedin 4.5. If $\mu(.|x) << \gamma$ for all $x\varepsilon X$ then $N(S^*, \lambda, U) = N(S, \lambda, U)$ for all $0 < \lambda < 1$.

Proof:

Suppose there exists $0 < \lambda < 1$ such that $N(S, \lambda, U) = 1$. Given $x_o\varepsilon X$ then, since $\mu << \gamma$, there exists a connected set $D\varepsilon Y'$ such that $\mu(D|x_o) = 1-\lambda$. Hence $\mu(D^c|x_o) = \lambda > 0$. Let $x'\varepsilon X$ be such that $\mu(D|x)$ has a maximum at x'. Then $x'\varepsilon X_\lambda$.

The following example demonstrates a channel with an uncountable input alphabet, in fact a closed interval, where the set of admissible measures is countable.

4.7 Example: Let $X = [o, 1]$ and let (Y, Y') be the real numbers and the Borel sets respectively. Define measures μ_1 and μ_2 by

$$\mu_1(A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-y^2/2}\, dy,$$

$$\mu_2(A) = \frac{1}{2\sqrt{2\pi}} \int_A e^{-y^2/8}\, dy$$

for each $A\epsilon Y'$. Let $X_2 = \{x|x\epsilon X,\ x \text{ is irrational}\}$; let $X_1 = X - X_2$. The family of measures $\{\mu(.|x)|x\epsilon X\}$ is defined by

$$\mu(A|x) = \mu_1(A-x) \text{ if } x\epsilon X_1 - \{o\} - \{1\}$$

$$\mu(A|x) = \mu_2(A-x) \text{ if } x\epsilon X_2 \cup \{o\} \cup \{1\}$$

for each $A\epsilon Y'$.

It is easy to see that $X_1 \subset X^*$. In fact, if $x = 0$ let $A = (-\infty, o)$; if $x = 1$ let $A = (o, \infty)$; if $x\epsilon X_1 \cap (o, 1)$ let $A = (x-1, x+1)$. Then $\mu(A|x) > 0$, $\mu(A^c|x) > 0$ and $\mu(A|x) \geq \mu(A|x')$ for all $x'\epsilon X$.

Consider $x_2\epsilon X_2$. Let $A\epsilon Y'$, A connected, with $\mu(A|x) > 0$ and $\mu(A^c|x') > 0$ for some $x'\epsilon X$. Then there exists α, $\beta\epsilon R$

with $A = (x-\alpha, x+\beta)$ and either $|x-\alpha| < \infty$ or $|x+\beta| < \infty$. If either $x-\alpha \in (0, 1)$ or $x+\beta \in (0, 1)$ there exists $x_1 \in X_1 \cap (0, 1)$ such that $\mu(A|x_1) > \mu(A|x_2)$. If $x+\beta \leq 0$ then $\mu(A|0) > \mu(A|x_2)$, and if $x-\alpha \geq 1$ then $\mu(A|1) > \mu(A|x_2)$. It follows that $x_2 \notin X^*$.

CHAPTER V

SUFFICIENT CONDITIONS FOR FINITE CAPACITY

Since many of the studies in information theory involve the channel capacity, it is highly desirable to know when the capacity is finite. In this chapter, sufficient conditions are obtained for a channel with additive noise to have finite capacity. It will be assumed that the input and output are subsets of the real numbers.

5.1 Lemma: Let $M = \{\mu(.|x) | x \in X$ be a family of probability measures defined on a measure space (Y, Y'). Suppose there exists a probability measure γ such that $\mu(.|x) \ll \gamma$ for all $x \in X$ and $\{d\mu(.|x)/d\gamma | x \in X\}$ is uniformly bounded. Given $n > o$ let $\mu^n(.|u)$ and γ^n represent the product measures an $(Y^n, (Y^n)')$. Then given $\epsilon > o$ there exists $\delta > o$ such that given $u \in X^n$ and $A \in (Y^n)'$ then $\mu^n(A|u) \geq \epsilon$ implies $\gamma^n(A) \geq \delta^n$.

Proof:

Let $f(y|x) = d\mu(.|x)/d\gamma$. Suppose $\{f(y|x) | x \in X\}$ is uniformly bounded by $M \geq 1$. Let $\epsilon > o$ be given. Let $\delta = \epsilon/M$. Given $u \in X^n$ let $f^n(y|u) = d\mu^n(.|u)/d\gamma^n$. Then $f^n(y|u) \leq M^n$. Hence if $\mu^n(A|u) \geq \epsilon$ then $\gamma^n(A) \geq \epsilon/M^n \geq (\epsilon/M)^n = \delta^n$.

The preceding lemma enables one to determine $N(S^n, \lambda)$ for any fixed values of n and $\lambda > 0$ provided that there exists a probability measure γ with respect to which the family $\{d\mu(.|x)/d\gamma\}$ is uniformly bounded.

5.2 Theorem: Let $S = \{X, \mu(.|x), (Y, Y')\}$ be a channel. Suppose there exists a probability measured γ defined on (Y, Y') such that the family $\{d\mu(.|x)/d\gamma | x \in X\}$ is uniformly bounded. Then, given $0 < \lambda < 1$, there exists $\delta_\lambda > 0$ such that $N(S^n, \lambda) \leq \frac{1}{\delta_\lambda^n}$ for any $n > 0$.

Proof:

Let $0 < \lambda < 1$ be given. By the lemma, there exists a δ_λ such that for any $n > 0$ and any $u \in X^n$ $\mu^n(A^{(n)}|u) \geq 1-\lambda$ implies $\gamma^n(A^{(n)}) \geq \delta_\lambda^n$. Since γ^n is a probability measure, it is clear that there are at most $\frac{1}{\delta_\lambda^n}$ disjoint subsets of Y^n of γ^n measure δ_λ^n. The conclusion is now clear.

The following result, due to Kemperman [5], will be used to show that the channel defined in theorem 5.2 has finite capacity.

5.3 Theorem. Let S be a channel with capacity C. For $0 < \lambda < 1$ define

$$\overline{C}(\lambda) = \overline{\lim_{n\to\infty}} \log \frac{1}{n} N(S^n, \lambda).$$

Then for each $0 < \lambda < 1$

$$C \leq \overline{C}(\lambda).$$

5.4 Corollary: Let C be the capacity of the channel described in theorem 5.2. Then $C < \infty$.

Proof:

Let $0 < \lambda < 1$ be given. By theorem 5.2, there exists $\delta_\lambda > 0$ such that $N(S^n, \lambda) \leq \frac{1}{\delta_\lambda^n}$. Hence,

$$C \leq \overline{C}(\lambda) \leq \overline{\lim_{n\to\infty}} \frac{1}{n} \log \left(\frac{1}{\delta_\lambda^n}\right) = -\log \delta_\lambda < \infty$$

5.5 Theorem: Let $S = \{X, \mu(.|x), (Y, Y')\}$ be a channel with additive noise μ which is absolutely continuous with respect to Lebesgue measure γ. If there exists a choice for the Radon-Nikodyn derivative $f = d\mu/d\gamma$ such that $\int g(y)d\gamma < \infty$ where $g(y) = \sup \{f(y|x) | x\epsilon X\}$ then the capacity of S is finite.

Proof:

Suppose there exists a choice for $d\mu/d\gamma$, say f, such that $\int g(y)d\gamma < \infty$. A totally finite measure, γ', is defined on (Y, Y') by $\gamma'(A) = \int_A g(y)d\gamma$ for all $A\epsilon Y'$.

Let $h(y|x) = d\mu(.|x)/d\gamma'$. Then, if $d\gamma'/d\gamma \neq 0$,

$h(y|x) = \frac{d\mu(.|x)}{d\gamma} \cdot \frac{d\gamma}{d\gamma'} = \frac{f(y|x)}{g(y)} \leq 1$. Hence $\{h(y|x) | x\epsilon X\}$ is uniformly bounded almost everywhere $[\gamma']$.

Define $\psi(A) = \gamma'(A)/\gamma'(Y)$ for all $A\epsilon Y'$. Then ψ is a probability measure and it is clear that $\{d\mu(.|x)/d\psi | x\epsilon X\}$ is uniformly bounded a.e. $[\psi]$. Thus by 5.4, the capacity of S is finite.

<u>5.6 Corollary</u>: If S is a channel of type A and f is a bounded bell function then the capacity of S is finite. In particular, if S has additive Gaussian Noise the capacity of S is finite.

The hypothesis of theorem 5.5 requires that there must exist a choice for $d\mu/d\gamma$ which is bounded. The foll wing example provides a partial justification for this restriction.

<u>5.7 Example</u>: Let S be the channel of type A defined by: $X = [0,1]$, Y is the real numbers, Y' is the Borel sets, and, given $A\epsilon Y'$ $\mu(A) = \int_{A_1} \frac{dx}{2x(-\ln x)^{3/2}}$ where $A_1 = A \cap [0, e^{-1}]$. It is clear that $\mu << \gamma$. In fact $f(x) = \frac{1}{2x(-\ln x)^{3/2}}$ almost everywhere $[\gamma]$ where f(x) is any choice for $d\mu/d\gamma$. It is easy to see that $\lim_{x\to 0} f(x) = \infty$. Let F(x) be the distribution function of μ. Then

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{1}{\sqrt{-\ln x}} & \text{if } 0 < x \leq e^{-1} \\ 1 & \text{if } x \geq e^{-1} \end{cases}$$

Hence, given $0 < \lambda < 1$, $\gamma(A) = e^{-(\frac{1}{1-\lambda})^2}$ for any set A of minimal γ-measure such that $\mu(A) = 1-\lambda$. It follows, since $X = [0, 1]$, that $N(S, \lambda) \geq e^{(\frac{1}{1-\lambda})^2}$. Therefore $\lim_{\lambda \to 1} (1-\lambda) \log N(S, \lambda) \geq \lim_{\lambda \to 1} \frac{1}{1-\lambda} = \infty$. It follows from Fano's theorem that the capacity of S is infinite.

<u>5.8 Remark</u>: Let $S = \{X, \mu(.|x), (Y, Y')\}$ be a channel. If X can be written in the form $X = \sum_{i=1}^{n} X_i$, $X_i \cap X_j = \emptyset$ for $i \neq j$, such that the hypothesis of theorem 5.5 is true for each of the subchannels $S_i = \{X_i, \mu(.|x), (Y, Y')\}$, then the capacity of C is finite. In theorem 5.5 it is shown that for each i there exists a probability measure ψ_i and $0 < m_i < \infty$ such that $d\mu(.|x)/d\psi_i \leq m_i$ for all $x \in X_i$. Let $\psi = \frac{1}{n} \sum_{i=1}^{n} \psi_i$. Then ψ is a probability measure and $d\mu(.|x)/d\psi \leq n \max \{m_i | 1 \leq i \leq n\}$ for all $x \in X$.

BIBLIOGRAPHY

1. Feinstein, A., "A New Basic Theorem of Information Theory," Trans IRE, PGIT, 2, (1954).

2. Fano, R. M., "Statistical Theory of Communication," Notes of a Course given at Mass. Inst. Techn., (1952, 1954).

3. Gnednko, B. V. and Kolomogorov, A. N., Limit Distributions of Independent Random Variables, Cambridge, Mass., Addison - Wesley Publishing Co., Inc., (1954).

4. Halmos, P. R., Measure Theory. New York, D. Van Nostrand Co., Inc., 1950.

5. Kemperman, J. H. B., "Studies in Coding Theory I," Unpublished Monograph, Univ. of Rochester.

6. Loeve, M., Probability Theory, 2nd. ed., New York, D. Van Nostrand Co., Inc., 1960.

7. Shannon, C. E., "A Mathematical Theory of Communication," Bell Systems Tech J., 21, 379-423, (1948).

8. ____________, "Certain Results in coding Theory for Noisy Channels," Inform. and Control, 1, 6-25, (1957).

9. Wolfowitz, J., "The Coding of Messages Subject to Chance Errors," Illinois J. Math, 1, No. 4, 591-606, (1957).

10. ____________, "Strong Converse of the Coding Theorem for Semi-Continuous Channels," Illinois J. Math., 3 No. 4, 477-489, (1959).

11. ____________, Coding Theorems of Information Theory, Berlin, Springer-Verlag, 1961.

ATMOSPHERIC SCIENCES RESEARCH PAPERS

1. Webb, W.L., "Development of Droplet Size Distributions in the Atmosphere," June 1954.
2. Hansen, F. V., and H. Rachele, "Wind Structure Analysis and Forecasting Methods for Rockets," June 1954.
3. Webb, W. L., "Net Electrification of Water Droplets at the Earth's Surface," *J. Meteorol.*, December 1954.
4. Mitchell, R., "The Determination of Non-Ballistic Projectile Trajectories," March 1955.
5. Webb, W. L., and A. McPike, "Sound Ranging Technique for Determining the Trajectory of Supersonic Missiles," #1, March 1955.
6. Mitchell, R., and W. L. Webb, "Electromagnetic Radiation through the Atmosphere," #1, April 1955.
7. Webb, W. L., A. McPike, and H. Thompson, "Sound Ranging Technique for Determining the Trajectory of Supersonic Missiles," #2, July 1955.
8. Barichivich, A., "Meteorological Effects on the Refractive Index and Curvature of Microwaves in the Atmosphere," August 1955.
9. Webb, W. L., A. McPike and H. Thompson, "Sound Ranging Technique for Determining the Trajectory of Supersonic Missiles," #3, September 1955.
10. Mitchell, R., "Notes on the Theory of Longitudinal Wave Motion in the Atmosphere," February 1956.
11. Webb, W. L., "Particulate Counts in Natural Clouds," *J. Meteorol.*, April 1956.
12. Webb, W. L., "Wind Effect on the Aerobee," #1, May 1956.
13. Rachele, H., and L. Anderson, "Wind Effect on the Aerobee," #2, August 1956.
14. Beyers, N., "Electromagnetic Radiation through the Atmosphere, #2. January 1957.
15. Hansen, F. V., "Wind Effect on the Aerobee," #3, January 1957.
16. Kershner, J., and H. Bear, "Wind Effect on the Aerobee," #4, January 1957.
17. Hoidale, G., "Electromagnetic Radiation through the Atmosphere," #3, February 1957.
18. Querfeld, C. W., "The Index of Refraction of the Atmosphere for 2.2 Micron Radiation," March 1957.
19. White, Lloyd, "Wind Effect on the Aerobee," #5, March 1957.
20. Kershner, J. G., "Development of a Method for Forecasting Component Ballistic Wind," August 1957.
21. Layton, Ivan, "Atmospheric Particle Size Distribution," December 1957.
22. Rachele, Henry and W. H. Hatch, "Wind Effect on the Aerobee," #6, February 1958.
23. Beyers, N. J., "Electromagnetic Radiation through the Atmosphere," #4, March 1958.
24. Prosser, Shirley J., "Electromagnetic Radiation through the Atmosphere," #5, April 1958.
25. Armendariz, M., and P. H. Taft, "Double Theodolite Ballistic Wind Computations," June 1958.
26. Jenkins, K. R. and W. L. Webb, "Rocket Wind Measurements," June 1958.
27. Jenkins, K. R., "Measurement of High Altitude Winds with Loki," July 1958.
28. Hoidale, G., "Electromagnetic Propagation through the Atmosphere," #6, February 1959.
29. McLardie, M., R. Helvey, and L. Traylor, "Low-Level Wind Profile Prediction Techniques," #1, June 1959.
30. Lamberth, Roy, "Gustiness at White Sands Missile Range," #1, May 1959.
31. Beyers, N. J., B. Hinds, and G. Hoidale, "Electromagnetic Propagation through the Atmosphere," #7, June 195[illegible]
32. Beyers, N. J., "Radar Refraction at Low Elevation Angles (U)," Proceedings of the Army Science Conference, June 1959.
33. White, L., O. W. Thiele and P. H. Taft, "Summary of Ballistic and Meteorological Support During IGY Operations at Fort Churchill, Canada," August 1959.
34. Hainline, D. A., "Drag Cord-Aerovane Equation Analysis for Computer Application," August 1959.
35. Hoidale, G. B., "Slope-Valley Wind at WSMR," October 1959.
36. Webb, W. L., and K. R. Jenkins, "High Altitude Wind Measurements," *J. Meteorol.*, 16, 5, October 1959.

37. White, Lloyd, "Wind Effect on the Aerobee," #9, October 1959.
38. Webb, W. L., J. W. Coffman, and G. Q. Clark, "A High Altitude Acoustic Sensing System," December 1959.
39. Webb, W. L., and K. R. Jenkins, "Application of Meteorological Rocket Systems," J. Geophys. Res., 64, 11, November 1959.
40. Duncan, Louis, "Wind Effect on the Aerobee," #10, February 1960.
41. Helvey, R. A., "Low-Level Wind Profile Prediction Techniques," #2, February 1960.
42. Webb, W. L., and K. R. Jenkins, "Rocket Sounding of High-Altitude Parameters," *Proc. GM Rel. Symp.*, Dept. of Defense, February 1960.
43. Armendariz, M., and H. H. Monahan, "A Comparison Between the Double Theodolite and Single-Theodolite Wind Measuring Systems," April 1960.
44. Jenkins, K. R., and P. H. Taft, "Weather Elements in the Tularosa Basin," July 1960.
45. Beyers, N. J., "Preliminary Radar Performance Data on Passive Rocket-Borne Wind Sensors," IRE TRANS, MIL ELECT, MIL-4, 2-3, April-July 1960.
46. Webb, W. L., and K. R. Jenkins, "Speed of Sound in the Stratosphere," June 1960.
47. Webb, W. L., K. R. Jenkins, and G. Q. Clark, "Rocket Sounding of High Atmosphere Meteorological Parameters," *IRE Trans. Mil. Elect.*, MIL-4, 2-3, April-July 1960.
48. Helvey, R. A., "Low-Level Wind Profile Prediction Techniques," #3, September 1960.
49. Beyers, N. J., and O. W. Thiele, "Meteorological Wind Sensors," August 1960.
50. Armijo, Larry, "Determination of Trajectories Using Range Data from Three Noncolinear Radar Stations," September 1960.
51. Carnes, Patsy Sue, "Temperature Variations in the First 200 Feet of the Atmosphere in an Arid Region," July 1961.
52. Springer, H. S., and R. O. Olsen, "Launch Noise Distribution of Nike-Zeus Missiles," July 1961.
53. Thiele, O. W., "Density and Pressure Profiles Derived from Meteorological Rocket Measurements," September 1961.
54. Diamond, M. and A. B. Gray, "Accuracy of Missile Sound Ranging," November 1961.
55. Lamberth, R. L. and D. R. Veith, "Variability of Surface Wind in Short Distances," #1, October 1961.
56. Swanson, R. N., "Low-Level Wind Measurements for Ballistic Missile Application," January 1962.
57. Lamberth, R. L. and J. H. Grace, "Gustiness at White Sands Missile Range," #2, January 1962.
58. Swanson, R. N. and M. M. Hoidale, "Low-Level Wind Profile Prediction Techniques," #4, January 1962.
59. Rachele, Henry, "Surface Wind Model for Unguided Rockets Using Spectrum and Cross Spectrum Techniques," January 1962.
60. Rachele, Henry, "Sound Propagation through a Windy Atmosphere," #2, February 1962.
61. Webb, W. L., and K. R. Jenkins, "Sonic Structure of the Mesosphere," *J. Acous. Soc. Amer.*, 34, 2, February 1962.
62. Tourin, M. H. and M. M. Hoidale, "Low-Level Turbulence Characteristics at White Sands Missile Range," April 1962.
63. Miers, Bruce T., "Mesospheric Wind Reversal over White Sands Missile Range," March 1962.
64. Fisher, E., R. Lee and H. Rachele, "Meteorological Effects on an Acoustic Wave within a Sound Ranging Array," May 1962.
65. Walter, E. L., "Six Variable Ballistic Model for a Rocket," June 1962.
66. Webb, W. L., "Detailed Acoustic Structure Above the Tropopause," *J. Applied Meteorol.*, 1, 2, June 1962.
67. Jenkins, K. R., "Empirical Comparisons of Meteorological Rocket Wind Sensors," *J. Appl. Meteor.*, June 1962.
68. Lamberth, Roy, "Wind Variability Estimates as a Function of Sampling Interval," July 1962.
69. Rachele, Henry, "Surface Wind Sampling Periods for Unguided Rocket Impact Prediction," July 1962.
70. Traylor, Larry, "Coriolis Effects on the Aerobee-Hi Sounding Rocket," August 1962.
71. McCoy, J., and G. Q. Clark, "Meteorological Rocket Thermometry," August 1962.
72. Rachele, Henry, "Real-Time Prelaunch Impact Prediction System," August 1962.

73. Beyers, N. J., O. W. Thiele, and N. K. Wagner, "Performance Characteristics of Meteorlogical Rocket Wind and Temperature Sensors," October 1962.
74. Coffman, J., and R. Price, "Some Errors Associated with Acoustical Wind Measurements through a Layer," October 1962.
75. Armendariz, M., E. Fisher, and J. Serna, "Wind Shear in the Jet Stream at WSMR," November 1962.
76. Armendariz, M., F. Hansen, and S. Carnes, "Wind Variability and its Effect on Rocket Impact Prediction," January 1963.
77. Querfeld, C., and Wayne Yunker, "Pure Rotational Spectrum of Water Vapor, I: Table of Line Parameters," February 1963.
78. Webb, W. L., "Acoustic Component of Turbulence," *J. Applied Meteorol.*, 2, 2, April 1963.
79. Beyers, N. and L. Engberg, "Seasonal Variability in the Upper Atmosphere," May 1963.
80. Williamson, L. E., "Atmospheric Acoustic Structure of the Sub-polar Fall," May 1963.
81. Lamberth, Roy and D. Veith, "Upper Wind Correlations in Southwestern United States," June 1963.
82. Sandlin, E., "An analysis of Wind Shear Differences as Measured by AN/FPS-16 Radar and AN/GMD-1B Rawinsonde," August 1963.
83. Diamond, M. and R. P. Lee, "Statistical Data on Atmospheric Design Properties Above 30 km," August 1963.
84. Thiele, O. W., "Mesospheric Density Variability Based on Recent Meteorological Rocket Measurements," *J. Applied Meteorol.*, 2, 5, October 1963.
85. Diamond, M., and O. Essenwanger, "Statistical Data on Atmospheric Design Properties to 30 km," *Astro. Aero. Engr.*, December 1963.
86. Hansen, F. V., "Turbulence Characteristics of the First 62 Meters of the Atmosphere," December 1963.
87. Morris, J. E., and B. T. Miers, "Circulation Disturbances Between 25 and 70 kilometers Associated with the Sudden Warming of 1963," *J. of Geophys. Res.*, January 1964.
88. Thiele, O. W., "Some Observed Short Term and Diurnal Variations of Stratospheric Density Above 30 km," January 1964.
89. Sandlin, R. E., Jr. and E. Armijo, "An Analysis of AN/FPS-16 Radar and AN/GMD-1B Rawinsonde Data Differences," January 1964.
90. Miers, B. T., and N. J. Beyers, "Rocketsonde Wind and Temperature Measurements Between 30 and 70 km for Selected Stations," *J. Applied Meteorol.*, February 1964.
91. Webb, W. L., "The Dynamic Stratosphere," *Astronautics and Aerospace Engineering*, March 1964.
92. Low, R. D. H., "Acoustic Measurements of Wind through a Layer," March 1964.
93. Diamond. M., "Cross Wind Effect on Sound Propagation," *J. Applied Meteorol.*, April 1964.
94. Lee, R. P., "Acoustic Ray Tracing," April 1964.
95. Reynolds, R. D., "Investigation of the Effect of Lapse Rate on Balloon Ascent Rate," May 1964.
96. Webb, W. L., "Scale of Stratospheric Detail Structure," *Space Research V*, May 1964.
97. Barber, T. L., "Proposed X-Ray-Infrared Method for Identification of Atmospheric Mineral Dust," June 1964.
98. Thiele, O. W., "Ballistic Procedures for Unguided Rocket Studies of Nuclear Environments (U)," Proceedings of the Army Science Conference, June 1964.
99. Horn, J. D., and E. J. Trawle, "Orographic Effects on Wind Variability," July 1964.
100. Hoidale, G., C. Querfeld, T. Hall, and R. Mireles, "Spectral Transmissivity of the Earth's Atmosphere in the 250 to 500 Wave Number Interval," #1, September 1964.
101. Duncan, L. D., R. Ensey, and B. Engebos, "Athena Launch Angle Determination," September 1964.
102. Thiele, O. W., "Feasibility Experiment for Measuring Atmospheric Density Through the Altitude Range of 60 to 100 KM Over White Sands Missile Range," October 1964.
103. Duncan, L. D., and R. Ensey, "Six-Degree-of-Freedom Digital Simulation Model for Unguided, Fin-Stabilized Rockets," November 1964.

104. Hoidale, G., C. Querfeld, T. Hall, and R. Mireles, "Spectral Transmissivity of the Earth's Atmosphere in the 250 to 500 Wave Number Interval," #2, November 1964.

105. Webb, W. L., "Stratospheric Solar Response," *J. Atmos. Sci.*, November 1964.

106. McCoy, J. and G. Clark, "Rocketsonde Measurement of Stratospheric Temperature," December 1964.

107. Farone, W. A., "Electromagnetic Scattering from Radially Inhomogeneous Spheres as Applied to the Problem of Clear Atmosphere Radar Echoes," December 1964.

108. Farone, W. A., "The Effect of the Solid Angle of Illumination or Observation on the Color Spectra of 'White Light' Scattered by Cylinders," January 1965.

109. Williamson, L. E., "Seasonal and Regional Characteristics of Acoustic Atmospheres," *J. Geophys. Res.*, January 1965.

110. Armendariz, M., "Ballistic Wind Variability at Green River, Utah," January 1965.

111. Low, R. D. H., "Sound Speed Variability Due to Atmospheric Composition," January 1965.

112. Querfeld, C. W., "Mie Atmospheric Optics," *J. Opt. Soc. Amer.*, January 1965.

113. Coffman, J., "A Measurement of the Effect of Atmospheric Turbulence on the Coherent Properties of a Sound Wave," January 1965.

114. Rachele, H., and D. Veith, "Surface Wind Sampling for Unguided Rocket Impact Prediction," January 1965.

115. Ballard, H., and M. Izquierdo, "Reduction of Microphone Wind Noise by the Generation of a Proper Turbulent Flow," February 1965.

116. Mireles, R., "An Algorithm for Computing Half Widths of Overlapping Lines on Experimental Spectra," February 1965.

117. Richart, H., "Inaccuracies of the Single-Theodolite Wind Measuring System in Ballistic Application," February 1965.

118. D'Arcy, M., "Theoretical and Practical Study of Aerobee-150 Ballistics," March 1965.

119. McCoy, J., "Improved Method for the Reduction of Rocketsonde Temperature Data," March 1965.

120. Mireles, R., "Uniqueness Theorem in Inverse Electromagnetic Cylindrical Scattering," April 1965.

121. Coffman, J., "The Focusing of Sound Propagating Vertically in a Horizontally Stratified Medium," April 1965.

122. Farone, W. A., and C. Querfeld, "Electromagnetic Scattering from an Infinite Circular Cylinder at Oblique Incidence," April 1965.

123. Rachele, H., "Sound Propagation through a Windy Atmosphere," April 1965.

124. Miers, B., "Upper Stratospheric Circulation over Ascension Island," April 1965.

125. Rider, L., and M. Armendariz, " A Comparison of Pibal and Tower Wind Measurements," April 1965.

126. Hoidale, G. B., "Meteorological Conditions Allowing a Rare Observation of 24 Micron Solar Radiation Near Sea Level," *Meteorol. Magazine*, May 1965.

127. Beyers, N. J., and B. T. Miers, "Diurnal Temperature Change in the Atmosphere Between 30 and 60 km over White Sands Missile Range," *J. Atmos. Sci.*, May 1965.

128. Querfeld, C., and W. A. Farone, "Tables of the Mie Forward Lobe," May 1965.

129. Farone, W. A., Generalization of Rayleigh-Gans Scattering from Radially Inhomogeneous Spheres," *J. Opt. Soc. Amer.*, June 1965.

130. Diamond, M., "Note on Mesospheric Winds Above White Sands Missile Range," *J. Applied Meteorol.*, June 1965.

131. Clark, G. Q., and J. G. McCoy, "Measurement of Stratospheric Temperature," *J. Applied Meteorol.*, June 1965.

132. Hall, T., G. Hoidale, R. Mireles, and C. Querfeld, "Spectral Transmissivity of the Earth's Atmosphere in the 250 to 500 Wave Number Interval," #3, July 1965.

133. McCoy, J., and C. Tate, "The Delta-T Meteorological Rocket Payload," June 1964.

134. Horn, J. D., "Obstacle Influence in a Wind Tunnel," July 1965.

135. McCoy, J., "An AC Probe for the Measurement of Electron Density and Collision Frequency in the Lower Ionosphere," July 1965.

136. Miers, B. T., M. D. Kays, O. W. Thiele and E. M. Newby, "Investigation of Short Term Variations of Several Atmospheric Parameters Above 30 KM," July 1965.

137. Serna, J., "An Acoustic Ray Tracing Method for Digital Computation," September 1965.

138. Webb, W. L., "Morphology of Noctilucent Clouds," *J. Geophys. Res.*, 70, 18, 4463-4475, September 1965.

139. Kays, M., and R. A. Craig, "On the Order of Magnitude of Large-Scale Vertical Motions in the Upper Stratosphere," *J. Geophys. Res.*, 70, 18, 4453-4462, September 1965.

140. Rider, L., "Low-Level Jet at White Sands Missile Range," September 1965.

141. Lamberth, R. L., R. Reynolds, and Morton Wurtele, "The Mountain Lee Wave at White Sands Missile Range," *Bull. Amer. Meteorol. Soc.*, 46, 10, October 1965.

142. Reynolds, R. and R. L. Lamberth, "Ambient Temperature Measurements from Radiosondes Flown on Constant-Level Balloons," October 1965.

143. McCluney, E., "Theoretical Trajectory Performance of the Five-Inch Gun Probe System," October 1965.

144. Pena, R. and M. Diamond, "Atmospheric Sound Propagation near the Earth's Surface," October 1965.

145. Mason, J. B., "A Study of the Feasibility of Using Radar Chaff For Stratospheric Temperature Measurements," November 1965.

146. Diamond, M., and R. P. Lee, "Long-Range Atmospheric Sound Propagation," *J. Geophys. Res.*, 70, 22, November 1965.

147. Lamberth, R. L., "On the Measurement of Dust Devil Parameters," November 1965.

148. Hansen, F. V., and P. S. Hansen, "Formation of an Internal Boundary over Heterogeneous Terrain," November 1965.

149. Webb, W. L., "Mechanics of Stratospheric Seasonal Reversals," November 1965.

150. U. S. Army Electronics R & D Activity, "U. S. Army Participation in the Meteorological Rocket Network," January 1966.

151. Rider, L. J., and M. Armendariz, "Low-Level Jet Winds at Green River, Utah," February 1966.

152. Webb, W. L., "Diurnal Variations in the Stratospheric Circulation," February 1966.

153. Beyers, N. J., B. T. Miers, and R. J. Reed, "Diurnal Tidal Motions near the Stratopause During 48 Hours at WSMR," February 1966.

154. Webb, W. L., "The Stratospheric Tidal Jet," February 1966.

155. Hall, J. T., "Focal Properties of a Plane Grating in a Convergent Beam," February 1966.

156. Duncan, L. D., and Henry Rachele, "Real-Time Meteorological System for Firing of Unguided Rockets," February 1966.

157. Kays, M. D., "A Note on the Comparison of Rocket and Estimated Geostrophic Winds at the 10-mb Level," *J. Appl. Meteor.*, February 1966.

158. Rider, L., and M. Armendariz, " A Comparison of Pibal and Tower Wind Measurements," *J. Appl. Meteor.*, 5, February 1966.

159. Duncan, L. D., "Coordinate Transformations in Trajectory Simulations," February 1966.

160. Williamson, L. E., "Gun-Launched Vertical Probes at White Sands Missile Range," February 1966.

161. Randhawa, J. S., Ozone Measurements with Rocket-Borne Ozonesondes," March 1966.

162. Armendariz, Manuel, and Laurence J. Rider, "Wind Shear for Small Thickness Layers," March 1966.

163. Low, R. D. H., "Continuous Determination of the Average Sound Velocity over an Arbitrary Path," March 1966.

164. Hansen, Frank V., "Richardson Number Tables for the Surface Boundary Layer," March 1966.

165. Cochran, V. C., E. M. D'Arcy, and Florencio Ramirez, "Digital Computer Program for Five-Degree-of-Freedom Trajectory," March 1966.

166. Thiele, O. W., and N. J. Beyers, "Comparison of Rocketsonde and Radiosonde Temperatures and a Verification of Computed Rocketsonde Pressure and Density," April 1966.

167. Thiele, O. W., "Observed Diurnal Oscillations of Pressure and Density in the Upper Stratosphere and Lower Mesosphere," April 1966.

168. Kays, M. D., and R. A. Craig, "On the Order of Magnitude of Large-Scale Vertical Motions in the Upper Stratosphere," *J. Geophy. Res.*, April 1966.

169. Hansen, F. V., "The Richardson Number in the Planetary Boundary Layer," May 1966.

170. Ballard, H. N., "The Measurement of Temperature in the Stratosphere and Mesosphere," June 1966.
171. Hansen, Frank V., "The Ratio of the Exchange Coefficients for Heat and Momentum in a Homogeneous, Thermally Stratified Atmosphere," June 1966.
172. Hansen, Frank V., "Comparison of Nine Profile Models for the Diabatic Boundary Layer," June 1966.
173. Rachele, Henry, "A Sound-Ranging Technique for Locating Supersonic Missiles," May 1966.
174. Farone, W. A., and C. W. Querfeld, "Electromagnetic Scattering from Inhomogeneous Infinite Cylinders at Oblique Incidence," *J. Opt. Soc. Amer.* 56, 4, 476-480, April 1966.
175. Mireles, Ramon, "Determination of Parameters in Absorption Spectra by Numerical Minimization Techniques," *J. Opt. Soc. Amer.* 56, 5, 644-647, May 1966.
176. Reynolds, R., and R. L. Lamberth, "Ambient Temperature Measurements from Radiosondes Flown on Constant-Level Balloons," *J. Appl. Meteorol.*, 5, 3, 304-307, June 1966.
177. Hall, James T., "Focal Properties of a Plane Grating in a Convergent Beam," *Appl. Opt.*, 5, 1051, June 1966
178. Rider, Laurence J., "Low-Level Jet at White Sands Missile Range," *J. Appl. Meteorol.*, 5, 3, 283-287, June 1966.
179. McCluney, Eugene, "Projectile Dispersion as Caused by Barrel Displacement in the 5-Inch Gun Probe System," July 1966.
180. Armendariz, Manuel, and Laurence J. Rider, "Wind Shear Calculations for Small Shear Layers," June 1966.
181. Lamberth, Roy L., and Manuel Armendariz, "Upper Wind Correlations in the Central Rocky Mountains," June 1966.
182. Hansen, Frank V., and Virgil D. Lang, "The Wind Regime in the First 62 Meters of the Atmosphere," June 1966.
183. Randhawa, Jagir S., "Rocket-Borne Ozonesonde," July 1966.
184. Rachele, Henry, and L. D. Duncan, "The Desirability of Using a Fast Sampling Rate for Computing Wind Velocity from Pilot-Balloon Data," July 1966.
185. Hinds, B. D., and R. G. Pappas, "A Comparison of Three Methods for the Correction of Radar Elevation Angle Refraction Errors," August 1966.
186. Riedmuller, G. F., and T. L. Barber, "A Mineral Transition in Atmospheric Dust Transport," August 1966.
187. Hall, J. T., C. W. Querfeld, and G. B. Hoidale, "Spectral Transmissivity of the Earth's Atmosphere in the 250 to 500 Wave Number Interval," Part IV (Final), July 1966.
188. Duncan, L. D. and B. F. Engebos, "Techniques for Computing Launcher Settings for Unguided Rockets," September 1966.
189. Duncan, L. D., "Basic Considerations in the Development of an Unguided Rocket Trajectory Simulation Model," September 1966.
190. Miller, Walter B., "Consideration of Some Problems in Curve Fitting," September 1966.
191. Cermak, J. E., and J. D. Horn, "The Tower Shadow Effect," August 1966.
192. Webb, W. L., "Stratospheric Circulation Response to a Solar Eclipse," October 1966.
193. Kennedy, Bruce, "Muzzle Velocity Measurement," October 1966.
194. Traylor, Larry E., "A Refinement Technique for Unguided Rocket Drag Coefficients," October 1966
195. Nusbaum, Henry, "A Reagent for the Simultaneous Microscope Determination of Quartz and Halides," October 1966.
196. Kays, Marvin and R. O. Olsen, "Improved Rocketsonde Parachute-derived Wind Profiles," October 1966.
197. Engebos, Bernard F. and Duncan, Louis D., "A Nomogram for Field Determination of Launcher Angles for Unguided Rockets," October 1966.
198. Webb, W. L., "Midlatitude Clouds in the Upper Atmosphere," November 1966.
199. Hansen, Frank V., "The Lateral Intensity of Turbulence as a Function of Stability," November 1966.
200. Rider, L. J. and M. Armendariz, "Differences of Tower and Pibal Wind Profiles," November 1966.
201. Lee, Robert P., "A Comparison of Eight Mathematical Models for Atmospheric Acoustical Ray Tracing," November 1966.
202. Low, R. D. H., et al., "Acoustical and Meteorological Data Report SOTRAN I and II," November 1966.

203. Hunt, J. A. and J. D. Horn, "Drag Plate Balance," December 1966.
204. Armendariz, M., and H. Rachele, "Determination of a Representative Wind Profile from Balloon Data," December 1966.
205. Hansen, Frank V., "The Aerodynamic Roughness of the Complex Terrain of White Sands Missile Range," January 1967.
206. Morris, James E., "Wind Measurements in the Subpolar Mesopause Region," January 1967.
207. Hall, James T., "Attenuation of Millimeter Wavelength Radiation by Gaseous Water," January 1967.
208. Thiele, O. W., and N. J. Beyers, "Upper Atmosphere Pressure Measurements With Thermal Conductivity Gauges," January 1967.
209. Armendariz, M., and H. Rachele, "Determination of a Representative Wind Profile from Balloon Data," January 1967.
210. Hansen, F. V., "The Aerodynamic Roughness of the Complex Terrain of White Sands Missile Range, New Mexico," January 1967.
211. D'Arcy, Edward M., "Some Applications of Wind to Unguided Rocket Impact Prediction," March 1967.
212. Kennedy, Bruce, "Operation Manual for Stratosphere Temperature Sonde," March 1967.
213. Hoidale, G. B., S. M. Smith, A. J. Blanco, and T. L. Barber, "A Study of Atmospheric Dust," March 1967.
214. Longyear, J. Q., "An Algorithm for Obtaining Solutions to Laplace's Tidal Equations," March 1967.
215. Rider, L. J., "A Comparison of Pibal with Raob and Rawin Wind Measurements," April 1967.
216. Breeland, A. H., and R. S. Bonner, "Results of Tests Involving Hemispherical Wind Screens in the Reduction of Wind Noise," April 1967.
217. Webb, Willis L., and Max C. Bolen, "The D-region Fair-Weather Electric Field," April 1967.
218. Kubinski, Stanley F., "A Comparative Evaluation of the Automatic Tracking Pilot-Balloon Wind Measuring System," April 1967.
219. Miller, Walter B., and Henry Rachele, "On Nonparametric Testing of the Nature of Certain Time Series," April 1967.
220. Hansen, Frank V., "Spacial and Temporal Distribution of the Gradient Richardson Number in the Surface and Planetary Layers," May 1967.
221. Randhawa, Jagir S., "Diurnal Variation of Ozone at High Altitudes," May 1967.
222. Ballard, Harold N., "A Review of Seven Papers Concerning the Measurement of Temperature in the Stratosphere and Mesosphere," May 1967.
223. Williams, Ben H., "Synoptic Analyses of the Upper Stratospheric Circulation During the Late Winter Storm Period of 1966," May 1967.
224. Horn, J. D., and J. A. Hunt, "System Design for the Atmospheric Sciences Office Wind Research Facility," May 1967.
225. Miller, Walter B., and Henry Rachele, "Dynamic Evaluation of Radar and Photo Tracking Systems, " May 1967.
226. Bonner, Robert S., and Ralph H. Rohwer, "Acoustical and Meteorological Data Report - SOTRAN III and IV," May 1967.
227. Rider, L. J., "On Time Variability of Wind at White Sands Missile Range, New Mexico," June 1967.
228. Randhawa, Jagir S., "Mesospheric Ozone Measurements During a Solar Eclipse," June 1967.
229. Beyers, N. J., and B. T. Miers, "A Tidal Experiment in the Equatorial Stratosphere over Ascension Island (8S)", June 1967.
230. Miller, W. B., and H. Rachele, "On the Behavior of Derivative Processes," June 1967
231. Walters, Randall K., "Numerical Integration Methods for Ballistic Rocket Trajectory Simulation Programs," June 1967.
232. Hansen, Frank V., "A Diabatic Surface Boundary Layer Model," July 1967.
233. Butler, Ralph L., and James K. Hall, "Comparison of Two Wind Measuring Systems with the Contraves Photo-Theodolite," July 1967.
234. Webb, Willis L., "The Source of Atmospheric Electrification," June 1967.

235. Hinds, B. D., "Radar Tracking Anomalies over an Arid Interior Basin," August 1967.
236. Christian, Larry O., "Radar Cross Sections for Totally Reflecting Spheres," August 1967.
237. D'Arcy, Edward M., "Theoretical Dispersion Analysis of the Aerobee 350," August 1967.
238. Anon., "Technical Data Package for Rocket-Borne Temperature Sensor," August 1967.
239. Glass, Roy I., Roy L. Lamberth, and Ralph D. Reynolds, "A High Resolution Continuous Pressure Sensor Modification for Radiosondes," August 1967.
240. Low, Richard D. H., "Acoustic Measurement of Supersaturation in a Warm Cloud," August 1967.
241. Rubio, Roberto, and Harold N. Ballard, "Time Response and Aerodynamic Heating of Atmospheric Temperature Sensing Elements," August 1967.
242. Seagraves, Mary Ann B., "Theoretical Performance Characteristics and Wind Effects for the Aerobee 150," August 1967.
243. Duncan, Louis Dean, "Channel Capacity and Coding," August 1967.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) U. S. Army Electronics Command Fort Monmouth, New Jersey 07703	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP
3. REPORT TITLE CHANNEL CAPACITY AND CODING	
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)	
5. AUTHOR(S) (First name, middle initial, last name) Louis D. Duncan	
6. REPORT DATE August 1967	7a. TOTAL NO. OF PAGES 64 — 7b. NO. OF REFS 11
8a. CONTRACT OR GRANT NO. b. PROJECT NO. 1L013001A91A c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) ECOM - 5151 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
10. DISTRIBUTION STATEMENT Distribution of this report is unlimited	
11. SUPPLEMENTARY NOTES Dissertation for Degree of Doctor of Philosophy, New Mexico State University	12. SPONSORING MILITARY ACTIVITY Atmospheric Sciences Laboratory U. S. Army Electronics Command White Sands Missile Range, New Mexico

13. ABSTRACT

This report presents original results in information. An improvement of Fano's theorem is proved. A sufficient condition for finite capacity for certain channels with additive noise is proved. An immediate corollary shows that the channel with additive Gaussian noise has finite capacity.

A new technique for coding - the connected code - is defined and investigated.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

UNCLASSIFIED
Security Classification

UNCLASSIFIED
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
1. Information Theory 2. Finite Capacity 3. Connected Code 4. Fano's Theorem						

UNCLASSIFIED
Security Classification